



Vaasan yliopisto  
UNIVERSITY OF VAASA

MINHAZ UDDIN

# **Integrating Physics-Informed and Machine Learning Models for Indoor Temperature Prediction**

School of Technology and Innovations  
Master's Thesis  
Sustainable and Autonomous systems

Vaasa 2025

---

**UNIVERSITY OF VAASA****School of Technology and Innovations****Author:** MINHAZ UDDIN**Title of the thesis:** Integrating Physics-Informed and Machine Learning Models for Indoor Temperature Prediction.**Degree:** Master of Science in Computing Sciences**Degree Programme:** Sustainable and Autonomous Systems**Supervisor:** Petri Välisuo**Co-supervisor:** Amit Shukla**Year:** 2025**Pages:** 69

---

**ABSTRACT:** Nordic buildings face increasing demands for energy efficiency and consistent indoor thermal comfort, creating accurate indoor temperature predictions increasingly crucial for intelligent HVAC control. This research addresses this challenge by proposing a hybrid prediction framework that combines physics-informed learning with traditional data-driven models to improve long-term accuracy.

This study uses a multiple-source dataset consisting of building sensors and satellite-derived solar radiation measurements. It includes important physical variables like the inside temperature, the temperature of the supply air, the temperature outside, and the amount of solar radiation. Using this multi-source dataset, several traditional machine learning models were developed for comparison such as Extreme Gradient Boosting, Feedforward Neural Networks, Random Forest, and Long Short-Term Memory and the proposed Physics-Informed Neural Network model.

Findings across the three forecasting horizons (short term: 1 hour, medium term: 6 hours, and long term: 12 hours) show that each model has the strengths at different periods- LSTM performs best in short-term prediction (MAE: 0.18), XGBoost in medium term prediction (MAE: 0.33), and the PINN achieves the most accurate (MAE: 0.29) in long-term prediction through its embedded physical constraints. Overall, the findings highlight the benefits of data-driven and physics-informed methods, and it emphasizes the potential of hybrid modelling approaches for supporting energy efficient, comfort oriented HVAC control in public buildings.

---

**KEYWORDS:** Machine Learning, Physics-Informed Neural Networks, LSTM, XGBoost, FNN, Prediction model, Indoor Prediction, Differential equations model, Smart Buildings, Multi-Horizon Forecasting.

## Contents

1	Introduction	9
1.1	Background and Motivation	11
1.1.1	Importance of Indoor Temperature Comfort and Energy Efficiency	11
1.1.2	Data Driven Modeling for Indoor Temperature Prediction	12
1.1.3	Opportunities of Physics-Guided Modeling	13
1.2	Problem Statement and Research Questions	14
1.2.1	Research Questions	15
1.3	Research Objectives	16
1.4	Scope and Limitation	17
1.5	Thesis Structure Overview	18
2	Literature Review	20
2.1	Indoor Temperature Prediction and its overview	20
2.2	Machine Learning Model for Indoor temperature modeling	21
2.3	Physics Informed Neural Networks	21
2.4	Physics-Informed Neural Networks for fast full-field temperature prediction	23
2.5	Summary of Research Gaps	24
3	Methodology	25
3.1	Overview of Methodological Approach	25
3.2	Data Preprocessing	26
	Handling Missing Data:	26
	Timestamp Alignment:	26
	Feature selection:	27
	Data Normalization:	27
3.3	Dataset Characteristics and Limitations	28
3.4	Feature Engineering	30
3.4.1	Selection of relevant features	32
3.4.2	Target variable creation	32
3.4.3	Temporal Structure for Forecasting	33

3.5	Model Architecture	33
3.5.1	Feedforward Neural Network	34
3.5.2	Long Short-Term Memory	35
3.5.3	Extreme Gradient Boosting	35
3.5.4	Random Forest	36
3.6	Physics-Informed Neural Network Modeling	36
3.6.1	Governing Differential Equation	37
3.6.2	Learning Physics Parameters	38
3.6.3	Physics Loss Integration	39
3.6.4	Machine Learning Model Comparison	40
3.7	Evaluation Metrics and Performance Comparison	40
3.7.1	Mean Absolute Error	41
3.7.2	Root Mean Squared Error	41
3.7.3	Coefficient of Determination	42
4	Results and Findings	43
4.1	FNN Architecture Search and Performance	43
4.2	One-Hour Prediction Window	44
4.3	Six-Hours Prediction Window	46
4.4	Twelve-Hours Prediction Window	48
4.5	Physics-Based ODE Model Performance	50
4.6	Physics-Informed Neural Network Performance	52
4.7	XGBoost Model Results	54
4.8	Random Forest Model Performance	57
4.9	LSTM Model Results	59
4.10	Model Performance Table	61
5	Conclusions	63
6	Limitations	64
	References	66
	Appendices	69

**Appendix 1. Github Link**

## Figures

<b>Figure 1 Architecture of the Feedforward Neural Network used for Indoor Temperature Prediction.</b>	<b>34</b>
<b>Figure 2: Architecture of the Physics-Informed Neural Network (PINN) for indoor temperature prediction</b>	<b>40</b>
<b>Figure 3 : One Hour Prediction: FNN vs Actual</b>	<b>45</b>
<b>Figure 4: One Hour Prediction: ODE(Physics) vs Actual</b>	<b>45</b>
<b>Figure 5: One Hour Prediction: FNN vs ODE vs PINN</b>	<b>45</b>
<b>Figure 6: Six-Hour Forecast: FNN vs Actual</b>	<b>47</b>
<b>Figure 7: Six-Hour Forecast: ODE vs Actual</b>	<b>47</b>
<b>Figure 8: Six-Hour Forecast: FNN vs ODE vs PINN</b>	<b>47</b>
<b>Figure 9: Twelve-Hour Prediction: FNN vs Actual</b>	<b>48</b>
<b>Figure 10: Twelve-Hour Forecast: ODE vs Actual</b>	<b>49</b>
<b>Figure 11:Twelve-Hour Forecast: FNN vs ODE vs PINN</b>	<b>49</b>
<b>Figure 12: Physics Parameter Learning</b>	<b>50</b>
<b>Figure 13: Physics Parameter Loss Over Epochs</b>	<b>51</b>
<b>Figure 14: PINN Data Loss Over Epochs</b>	<b>52</b>
<b>Figure 15:PINN Physics Loss Over Epochs</b>	<b>53</b>
<b>Figure 16: Data Loss+ Physics Loss Over Epochs</b>	<b>53</b>
<b>Figure 17: XGBoost One-Hour: Prediction vs Actual</b>	<b>55</b>
<b>Figure 18 : XGBoost Six-Hour: Prediction vs Actual</b>	<b>55</b>
<b>Figure 19: XGBoost Twelve-Hour : Prediction vs Actual</b>	<b>56</b>
<b>Figure 20: One-Hour Forecast: Random Forest vs Actual</b>	<b>57</b>
<b>Figure 21: Six-Hour Forecast: Random Forest vs Actual</b>	<b>58</b>
<b>Figure 22:Twelve-Hour Prediction: Random Forest vs Actual</b>	<b>58</b>
<b>Figure 23:One-Hour LSTM Prediction vs Actual</b>	<b>60</b>
<b>Figure 24: Six-Hour Prediction Plot</b>	<b>60</b>
<b>Figure 25: Twelve-Hour LSTM Prediction Plot</b>	<b>61</b>

## Tables

<b>Table 1 : Summary Statistics of Input Variables and Justifying the Use of RobustScaler</b>	<b>31</b>
<b>Table 2: FNN Architecture Search Parameters Used in This Study</b>	<b>43</b>
<b>Table 3 :Top 20 FNN Architectures Based on Validation MAE</b>	<b>44</b>
<b>Table 4:Performance Metrics for the One Hour Prediction (FNN vs PINN)</b>	<b>46</b>
<b>Table 5:Performance Metrics for the Six Hour Prediction (FNN Vs PINN)</b>	<b>48</b>
<b>Table 6: Performance Metrics for the Twelve Hour Prediction (FNN Vs PINN)</b>	<b>50</b>
<b>Table 7: Physics Learnable Parameters Value</b>	<b>51</b>
<b>Table 8: PINN Architecture Parameter Used In This Study</b>	<b>52</b>
<b>Table 9: PINNs Learnable Parameters Value</b>	<b>54</b>
<b>Table 10: XGBoost Architecture Parameters Used In This Study</b>	<b>54</b>
<b>Table 11: XGBoost Prediction Performance</b>	<b>56</b>
<b>Table 12: Random Forest Parameters Used in This Study</b>	<b>57</b>
<b>Table 13: Random Forest Prediction Performance</b>	<b>59</b>
<b>Table 14: LSTM Architecture Parameters Used In This Study</b>	<b>59</b>
<b>Table 15:Combined Model Performance Table</b>	<b>62</b>

## Abbreviations

<b>CFD</b>	Computational Fluid Dynamics
<b>CNN</b>	Convolutional Neural Network
<b>CO<sub>2</sub></b>	Carbon Dioxide
<b>EU</b>	European Union
<b>FNN</b>	Feedforward Neural Network
<b>HVAC</b>	Heating, Ventilation, and Air Conditioning
<b>IoT</b>	Internet of Things
<b>LSTM</b>	Long Short-Term Memory
<b>ML</b>	Machine Learning
<b>MLP</b>	Multi-Layer Perceptron
<b>MAE</b>	Mean Absolute Error
<b>ODE</b>	Ordinary Differential Equation
<b>PINNs</b>	Physics-Informed Neural Networks
<b>RF</b>	Random Forest (Regressor)
<b>ReLU</b>	Rectified Linear Unit
<b>RMSE</b>	Root Mean Squared Error
<b>R<sup>2</sup></b>	Coefficient of Determination (R-squared)
<b>SVM</b>	Support Vector Machine
<b>XGBoost</b>	Extreme Gradient Boosting

## 1 Introduction

The accumulation of greenhouse gases, mainly CO<sub>2</sub>, is one of the main causes of climate change and global temperature increase. This means that we need to be more focused on considerable emission reductions. In response, the European Union has set targets to minimize emissions by 55% by 2030 and to achieve climate neutrality by 2050 (Loffa et al., 2025). In Nordic countries, heating homes mostly uses a lot of energy and reason is long winters and cold weather. Which makes it necessary to keep the indoor temperature constantly controlled. An accurate real-time model for predicting indoor temperature distribution is essential for both energy efficiency and thermal comfort (Ma et al., 2021; Park et al., 2021), particularly when a workplaces or localized zones requires controlling (Sun et al., 2020). As building became progressively digitized, there is an opportunity to use sensor-based data and advanced algorithms to develop predictive models that can forecast temperature changes and optimize heating, ventilation, and control central or non-central air conditioning operations.

Conventional modeling approaches, such as physics-based simulations that use thermodynamic principles, are often limited by their complexity, high computational cost, and dependence on precise building specifications (Lu et al., 2020). In recent times, data-driven methods, mainly machine learning (ML) models, have gained more attention due to their ability to model nonlinear relationships from sensor data without having complete physical knowledge (Amasyali & El-Gohary, 2018).

Among various ML methodologies, deep learning models such as Feedforward Neural Networks (FNN), Long Short-Term Memory networks (LSTM), and Extreme Gradient Boosting (XGBoost) have been utilized with different levels of efficacy for predicting indoor environment characteristics (Li et al., 2025; Li et al., 2020; Norouzi et al., 2023). However, a major limitation of purely data-driven models is their lack of physical interpretability and their inability to generalize under novel conditions. As a result, physics-informed neural networks (PINNs) have emerged. These help to make sure that

model predictions are within the limits of what science says is possible. They normally do this by adding physics rules and laws directly into the loss function during training period (Raissi et al., 2019). The most interesting part of PINNs is that it ensures that model predictions not only fit observed data but also obey critical physical principles.

According to Nordic climate where temperature fluctuations are significant, it is essential to accurately predict indoor temperature to maintain thermal comfort and reduce unnecessary heating energy use. Reliable prediction models support the proactive control of Heating, Ventilation, and Air Conditioning (HVAC) systems, and it prevents energy waste by ensuring comfortable indoor conditions for occupants. In the context of indoor temperature prediction, combining the physics ordinary differential equation (ODE) with ML models has shown promise in enhancing prediction accuracy while maintaining physical accuracy also (Hannula et al., 2025; Häkkinen, n.d.). This is a hybrid approach that aims to create a bridge between theoretical and data-driven modeling. It enables significant improvements to HVAC control strategies. Additionally, Internet of things (IoT)-enabled sensors and smart buildings have made huge amounts of sensor data available, which allows for more accurate and detailed modeling structure. According to the researcher (Miller et al., 2018), these datasets are kind of an ideal foundation for creating intelligent control systems that can dynamically react to user behavior, internal thermal loads, and other external variables. By using this data in hybrid modeling frameworks, it is possible to predict indoor temperature trends across various time periods and implement measures to prevent any issues.

This thesis investigates the importance of PINNs for indoor temperature prediction in public buildings, while comparing their performance with traditional ML models. The goal is to develop a reliable and generalizable model that can support the energy-efficient operation of buildings under any real-world conditions.

## **1.1 Background and Motivation**

### **1.1.1 Importance of Indoor Temperature Comfort and Energy Efficiency**

Especially in the Nordic region, indoor thermal comfort maintenance and maintaining energy efficiency are one of the main pillars in modern building operations. One of the main reasons behind it is the long winter season in Nordic countries, which makes consistent heating demand. As it is known, thermal comfort has major impacts on human health, cognitive performance, and productivity (Luo et al., 2023). Improper maintenance in indoor conditions can provide discomfort, as a result, it can reduce concentration, increase fatigue, and affect long-term health (Du et al., 2023). In research it is found that in the EU, its buildings are responsible for about 40% of the total energy consumption and 36% of CO<sub>2</sub> emissions, where space heating being is being found the primary energy consumption (Economidou et al., 2020). As a result, international and regional energy policies such as the EU green deal are focused on retrofitting and operational optimization of existing buildings and it is hope to meet climate neutrality goals by 2050 (Loffa et al., 2025).

Predictive modeling, like indoor temperature prediction or comfort level prediction, becomes an important part of a smart building control system. By predicting indoor temperature changes, operations of HVAC systems can be adjusted carefully, unnecessary energy use can be reduced, and thermal discomfort can be avoided also. The success of this kind of model depends on its ability to process complex relationships in the data and adapt to dynamic scenarios (Amasyali & El-Gohary, 2018).

Importantly, these predictive models can be improved and provide more accurate results by aligning them with physical law. As researchers (Hannula et al., 2025; Häkkinen, n.d.) demonstrated, that the models which mainly integrate physics-based constraints with ML, such as PINNs which has a chance that it can perform better in real-world

applications by ensuring physical consistency and improving generalization under variable conditions.

Hence, optimizing the indoor temperature is not just always about comfort, it is also important for reducing operational costs, achieving efficient energy goals and making perfect contribution to broader climate change mitigation.

### **1.1.2 Data Driven Modeling for Indoor Temperature Prediction**

Data-driven approaches, especially ML models, which have shown significant improvement and promised results in indoor temperature forecasting because it can understand and it has the ability to learn complex, nonlinear patterns from sensor data. Indoor temperature predictions and its changes is mostly done with ML models like XGBoost, LSTM, Convolutional Neural Network with Long Short-Term Memory (CNN-LSTM), and Multi-layer Perception (MLP) (Hannula et al., 2025; Li et al., 2020; Häkkinen, n.d.). Although, all these individual data-driven models face several challenges that constrain the accuracy, robustness, and generalization.

One significant challenge is dependence on better-quality representative training data. In the case of most building environments, sensor placement, calibration errors, and data gaps give noisy data and sometimes an incomplete dataset. This has negative effects on model training and reduce the prediction reliability also (Amasyali & El-Gohary, 2018; Miller et al., 2018). Moreover, these ML models are sensitive to changes in external environmental conditions such as weather changes, occupancy, humidity, and there is a chance that it can perform well in training environments but struggle when it will be deployed in an environment that is much different from their training data or unseen condition (Yan et al., 2015).

Another limitation is the “black box” nature of most ML models. Though this kind of model can provide accurate predictions in under some controlled scenarios, but it generally fails to explain how these kinds of predictions are made. It is therefore difficult

to ensure that the model always works, which would be needed for safety-critical for HVAC control application type situations (Li et al., 2025). Traditional ML models generally require large quantities of labelled data for training, which may not always be possible and feasible in many buildings, especially older buildings. As sensor installation is not feasible on those buildings, or it will be much more costly.

To address these limitations, recent research proposes that integrating physics domain knowledge with ML, such as PINNs, which mainly aims to embed thermodynamic principles into learning processes and uses that knowledge to train the model. So that the trained model is expected to predict more accurately. These kind of hybrid approaches can improve prediction accuracy and provide more sustainable solutions in intelligent building systems (Hannula et al., 2025; Raissi et al., 2019; Amasyali & El-Gohary, 2018).

### **1.1.3 Opportunities of Physics-Guided Modeling**

Physics-Informed modelling, offers some promising direction for the improvement of indoor temperature prediction in buildings. By introducing and integrating physics laws directly into the learning process of a model, it can overcome many weaknesses found in purely data-driven methods (Amasyali & El-Gohary, 2018; Raissi et al., 2019).

One of the main benefits of physics-informed approaches is their ability to use the thermodynamic principle for training the model or use it for physics learning. For example, while forecasting indoor temperature, the heat transfer equation based on Fourier law of heat conduction can be added, or the energy balance principle can also be integrated as a part of the physics function. This mainly allows the model not only to learn from available data but also helps to know physics behaviour as it is leading to more stable and realistic predictions (Hannula et al., 2025; Häkkinen, n.d.).

Another important opportunity is improved data efficiency. ML models mainly require large datasets to generalize well, which might not be always possible and feasible. However, this kind of assumption is not true, mainly in practical environments where limited amount of data is available. PINNs models can achieve better performance even with a small dataset, as it provides additional structure which can be provided into ML models and guide the learning process. This makes it a perfect fit for applications using historical or sensor data (Zhao et al., 2021).

Moreover, the hybrid design of PINN models provides sufficient flexibility to integrate ML architectures such as LSTM and FNN, which also has the ability of learning complex and nonlinear relationship patterns . This balance is especially useful for smart HVAC systems because trust in predictions is crucial for automated decision-making (Jaffal, 2023).

Recent research also shows that PINN models perform more accurately under some extreme and changing conditions. For instance, there is one research that explain that PINNs can provide better prediction accuracy in colder periods compared to standard LSTM model, mainly due to their alignment with thermodynamic behavior in the extrapolation scenario (Hannula et al., 2025; Häkkinen, n.d.).

In summary, integrating physics laws and their knowledge into ML creates opportunities to improve robustness, accuracy, and reliability in indoor temperature forecasting. This makes physics-guided modeling particularly gorgeous and sustainable for smart building control strategies.

## **1.2 Problem Statement and Research Questions**

The increasing demand for energy efficiency and indoor thermal comfort in different buildings of Nordic countries has given more priority in the development of accurate and reliable indoor temperature prediction models. While in traditional models and their learning shows promising results in complex modeling and thermal behavior, which

mainly uses sensor and some environmental data (Amasyali & El-Gohary, 2018; Jaffal, 2023). However, physics-informed models need more detailed input parameters. They are also kind of more computationally intensive (Zhao et al., 2021). As limited research has been conducted comparing PINNs with traditional models like LSTM, XGBoost, and FNN on building sensor datasets, this study aims to fill that gap by evaluating their performance using multi-source input data.

Indoor temperature can be predicted using combination of real time and historical data from various sources. These sources include building sensor data such as current indoor temperature, districts heating water temperature, humidity, and CO<sub>2</sub> levels, HVAC system inputs such as supply air temperature, return air temperature, air flow rate, district heating water incoming and outgoing temperature. Satellite based data like surface temperature and solar radiation can also add and support more environmental context to it.

Predicting indoor temperature is important because thermal systems can have some time delays means that by the time a change in temperature occurred and detected by a controller the comfortness of may already compromised. So predictive models enable a proactive suggestion to HVAC control system which allows the system to have some idea of future conditions such as sudden drop in outdoor temperature then it will adjust heating or cooling before any discomfort happens. So, this is not only improves thermal comfort but also optimize energy by avoiding over heating or under heating the indoor.

Therefore, in this thesis, both data-driven models and physics equation-based learning approach are used by developing a hybrid model for indoor temperature forecasting.

### **1.2.1 Research Questions**

- How accurately indoor temperature can be predicted using ML models and which model is the most suitable?

- How PINNs can improve the accuracy of indoor temperature prediction compared to traditional data driven models?
- Can multi-horizon prediction improve the reliability of indoor temperature predictions for HVAC control in smart buildings?

### **1.3 Research Objectives**

The primary objective of this thesis is to develop a prediction model that integrates physical knowledge in a form of ODE using outdoor temperature, solar radiation, ventilation and other data to model heat transfer. This approach enables the prediction of air temperature based on the range of features, added with a physics-consistent and generalizable indoor temperature prediction model for residential and public buildings. This research aims to bridge the gap between traditional ML models and physics-based learning and its approaches. PINNs can be used to connect domain knowledge to the learning process., we can achieve more accurate results. The study shows the effectiveness of this hybrid approach in comparison to baseline models such as LSTM, XGBoost, and FNN. The objectives of the study are articulated below:-

#### **Objective 1: To develop and evaluate traditional machine learning models (LSTM, XGboost, FNN, Random forest)**

This study aims to create a model using indoor sensor data, satellite based solar radiation input and other environmental parameters for indoor temperature prediction and evaluating the ML performance.

#### **Objective 2: To construct a physics-informed neural network**

To construct a model by embedding thermodynamic principles into it and provide loss functions like physics loss and data loss to guide the model's Learning process towards physically consistent predictions.

#### **Objective 3: To compare the predictive accuracy, generalization ability, and physical consistency**

Compare the models like XGboost, LSTM, Random forest which are traditional models, and the PINNs model across multiple forecasting horizons and check which one is working more efficiently for prediction.

#### 1.4 Scope and Limitation

This study aims to predict indoor temperatures in both residential and non-residential buildings in Nordic climate zones. Where heating plays an important role in overall energy consumption. The scope is defined by the use of real-world sensor data and historical data from systems. The models are trained and tested using data collected from smart buildings equipped with sensors from Finland (Nordic Zone).

The core of this study lies in the comparative evaluation of traditional ML models such as (FNN, LSTM, XGBoost) and PINN model which is also a combination of FNN model with combination of physics loss and data loss. The PINN is constructed by integrating a simplified thermodynamic ODE which represent indoor thermal dynamics and guiding the learning process toward physically meaningful predictions.

However, this study has several limitations:

- **Simplified physics representation:** The thermal behavior of buildings is influenced by multiple interacting factors such as solar gains, outdoor temperature, humidity, and indoor air temperature. The physics equation used in this thesis captures only this subset of dynamics.
- **Building-specific modeling:** The models are developed using data from a limited number of buildings. Therefore, generalization to buildings with different structures, insulation properties or usage patterns may be limited without retraining or adaptation.
- **Data quality and missing values:** Although with lots of preprocessing efforts missing or noisy sensor data can introduce an uncertainty in both data-driven and PINN models.

In spite of these limitations, the results teach us important things about the good and bad parts of using hybrid learning for smart building energy prediction and contribute to the emerging body of research on physics based learning in the built environment.

## **1.5 Thesis Structure Overview**

### **Chapter 1: Introduction**

This chapter introduces the motivation behind the study. It also outlines the research problem. The chapter identifies the existing research gap. It also presents the objectives and scope of the work. It also provides information about indoor temperature prediction and explains why combining physics-informed learning with ML models is important for improving energy efficiency in smart buildings.

### **Chapter 2: Literature Review**

This chapter reviews the existing structure of research related to indoor temperature prediction, energy-efficient building operation, and data-driven modeling techniques, like: FNN, LSTM, and XGBoost. It also discusses PINNs model as well. The chapter also highlights the limitations of purely data-driven methods and the opportunities presented by hybrid modeling approaches.

### **Chapter 3: Methodology**

In this chapter, it mainly describes the data sources, feature engineering, and preprocessing steps. Details of the development of baseline ML models and the physics-informed model. Also explains the physics-guided modeling components and governing equations and how they are integrated, and defines the evaluation metrics used for compare model performance.

### **Chapter 4: Results and Findings**

In this chapter, it presents the experimental results for all models and prediction horizons. The performance is compared based on MAE, RMSE, and  $R^2$ , and also the prediction behavior is shown using a time-series plots. The chapter highlights three major findings.

- 1) LSTM performs best in short-term 1 hour predictions.
- 2) XGBoost performs competitively for 6 hour forecasts.
- 3) PINN model provides the more stable prediction for 12 by integrating physical constraints.

Additionally, the chapter also analyzes error behavior, model limitations, and cases where models struggle to find sudden temperature fluctuations.

### **Chapter 5: Conclusions**

This chapter summarizes the significant scientific contributions and insights obtained through the comparative evaluation. It explains how hybrid physics-informed approaches improve the stability and PINN model accuracy of long-term predictions.

### **Chapter 6: Limitations**

This chapter highlights key limitations in both the practical and methodological aspects. These limitations include sensor uncertainty, sensor placement, sudden temperature spikes, and limited ground-truth reliability. The chapter also suggests directions for future research, including comfort-based prediction, and advanced PINN frameworks.

## 2 Literature Review

This chapter mainly identifies the existing research that has been done on prediction using data-driven and physics-informed ML models. The review highlights main themes such as conventional thermodynamic modeling, the rise of ML approaches like FNN, XGBoost, LSTM, and the recent emergence of PINNS. The key purpose of this review is to build a theoretical foundation that will guide through the hybrid modeling approach that is adopted in this thesis. This kind of theoretical foundation mainly supports the development of physically consistent, scalable, and interpretable predictive models that can be used for real-world building data and environmental data for predicting indoor climate control optimization.

### 2.1 Indoor Temperature Prediction and its overview

Indoor temperature prediction plays an important role in an efficient energy control system, like an HVAC system. Accurate prediction supports real-time decision making and demand side management. This is important mainly in buildings where maintaining thermal comfort is critical (Ma et al., 2021). Traditional physics-based simulations using EnergyPlus or Modelica are based on the principles of thermodynamics. Despite this, their deployment is often hampered by the need for detailed building metadata, long computational runtimes, and scalability issues (Luo et al., 2023; Miller et al., 2018).

In recent years data driven ML models such as LSTM, XGBoost, CNN can solve complex and nonlinear issues from historical sensor data which enables more accurate prediction without deep knowledge of building physics. Although this model may not perform well or have lack of interpretability in unseen scenarios (Häkkinen, n.d.; Hannula et al., 2025). So for this reason physics informed approaches is introduced as it can combine both ML with the structure of physics laws by integrating equations like furrier equation, energy balance equations improve its robustness (Zhao et al., 2021; Hannula et al., 2025; Häkkinen, n.d.).

## 2.2 Machine Learning Model for Indoor temperature modeling

Particularly in ML, data-driven ML models were able to gain more popularity for indoor temperature prediction as it has the ability to learn complex, nonlinear relationships from sensor data without requiring any detailed physics knowledge of building. The increasing availability of IoT based sensor data such as indoor temperature, outdoor condition, ventilation settings and solar radiation has further boosted the use of these methods in smart buildings (Amasyali & El-Gohary, 2018). Models such as FNN, Random Forests (RF), Support Vector Machines (SVM), and XGBoost have been widely used for indoor temperature forecasting (Amasyali & El-Gohary, 2018). FNNs and XGBoost are particularly suited for capturing nonlinearities and short-term dynamics, but they have one issue such as a lack of long-term temporal dependencies. To overcome this, LSTM architectures have been added in recent studies (Li et al., 2020). As of now, LSTMs can learn temporal dependencies more effectively, especially in hourly or sub-hourly building sensor datasets.

However, purely data-driven models tend to act as black box modeling. Their performance deteriorates when exposed to out-of-distribution inputs or sensor drifts, limiting their reliability for building automation systems (Loffa et al., 2025; Luo et al., 2023). Studies have shown that it often fail to generalize under changes in occupant behavior, external weather, or system faults (Muroli et al., 2019)

## 2.3 Physics Informed Neural Networks

Physics-Informed Neural Networks were mainly introduced as a hybrid modeling approach to address the shortcomings of ML models (Raissi et al., 2019). In PINNs, domain knowledge is embedded into the neural network training by incorporating a physics-based residual into the loss function. This gives a chance to the model to remain grounded in physical principles even when the data is sparse or noisy.

In the building energy modeling context, several studies have explored PINNs and physics-based learning. For instance, Ma et al. applied a physics-constrained deep learning model for indoor temperature forecasting and reporting improved generalization and consistency over standard LSTM (Ma et al., 2021). Hannula et al. developed a PINN using a simplified energy balance differential equation. Which demonstrates that it can outperform a regular LSTM in terms of Root Mean Squared Error (RMSE) and interpretability (Hannula et al., 2025). Similarly, Wang et al. extended this idea by integrating PINNs with sensor fusion techniques for better spatiotemporal temperature prediction (Wang et al., 2025).

A recent comparative study by titled Physics-Informed vs. Deep Learning Indoor Temperature, the authors (Loffa et al., 2025) mainly identifies the limitations of traditional ML and deep learning models, such as LSTM. The main reason for this modeling is forecasting indoor temperature under data-scarce or non-ideal conditions. It is recognized that data-driven models often act as black boxes with limited physical clarity. So, the authors propose a PINN framework that interconnects governing thermodynamic principles directly into its learning process. This approach emphasizes the integration of a residual-based loss function derived from the parameters of simplified ODEs. Two weeks to two years of data was used for prediction model training and for the evaluation of their performance with varying data volume. The dataset comprises simulated data that is generated using output from an Energy Plus simulation. For this research, Turin (Italy), Munich (Germany), Copenhagen (Denmark), and Madrid (Spain). These cities represent a diverse range of climates with hot summers and cold winters. Their findings show that the PINN models consistently outperform LSTM models in small data scenarios across multiple cities. In that research, with two weeks of training data, PINN achieved a lower mean absolute error (MAE) of 0.40 °C in Turin compared to LSTM, which is 0.56 °C. In the case of Munich, the PINNs also outperformed the LSTM across all periods, with the most stable performance, MAE of 0.15 °C (PINN) and 0.49 °C for (LSTM). Even in more complex environments like Madrid, the PINN maintains an error

below 2°C in all cases. The research highlights that incorporating physics laws into training enables PINNS to generalize better and provide reliable predictions.

## **2.4 Physics-Informed Neural Networks for fast full-field temperature prediction**

Accurately predicting indoor temperature is little challenging as it normally follows two approaches like physics-based models and data-driven based model. To overcome the limitations of both model (Jing et al., 2023) introduced a physics-informed framework of neural networks for predicting full field indoor temperatures. Their method combines CFD simulations, Physics laws, and sparse sensor data into a unified learning system. The framework has three components:

- A surrogate model that is trained on CFD data and ruled by energy conservation equations,
- A model that adjusts the surrogate model output using a small set of real sensor measurements through transfer learning.
- A recovery model that integrates both to produce fast and accurate temperature prediction across time series.

The authors tested these methods in a simulated room environment with changing air velocity and heat sources. With four observation points, the models performed well according to the author (RMSE = 0.777 °C,  $R^2 = 0.999$  at 600s). With more sensors (24 points), accuracy improved further (RMSE = 0.425 °C).

Although the training phase required significant CFD work and sensor data and once trained PINN enables real time accurate predictions for that environment. However, it has a limitation that it is not ideal for small building zones with complex heat sources due to modeling limitations.

## 2.5 Summary of Research Gaps

Existing studies (Jing et al., 2023; Loffa et al., 2025) and lots of other research mainly train and validate their frameworks on a simulated building or a limited set of climate zones. Their ability to generalize across the real sensor data of a building with data-driven model, physics-informed models and diverse weather patterns make a comparison between the models performance evaluation has remained in a gap. While many significant progresses have been made but gaps remain in model transferability, performance variation under noisy or non-noisy data, and integration with real-world HVAC control logic remains gap. Most current works evaluate models on isolated buildings without generalizing across zones with simulated data or varying weather scenarios or considering human comfort. Moreover, there are some works done with PINNs model, but there is less research done about the performance of PINNs model considering to others models, so we can say it is still in experimental stages.

### 3 Methodology

This chapter outlines the methodology for the development, training, and evaluation of ML models and hybrid physics-informed models for indoor temperature prediction. This research aims to provide a transparent and consistent process from data collection to model evaluation. Mainly, the workflow includes data preprocessing, model design, training strategy, performance evaluation, and then making a comparison between traditional ML models and physics-informed models. One of the main focuses of this research is the integration of physics-based constraints using the ODE of indoor heat transfer to enhance the model concept and physics consistency.

#### 3.1 Overview of Methodological Approach

The methodology that is followed in this thesis is a combination of both data-driven and physics-informed techniques to forecast indoor temperature with accuracy and consistency. Some key components of this methodology are:

- Collecting and preprocessing building sensor data and satellite based solar radiation observation data at 5-minute intervals.
- Feature engineering and identifying meaningful input variables.
- Model development using traditional models: FNN, XGBoost, and LSTM.
- Development of a hybrid PINN model, which is integrated with physics laws via ODEs.
- Multi-horizon or multi-interval prediction to evaluate short and long-term predicting efficiency.
- Evaluation of performance using performance metrics, including MAE and RMSE.

Each model was trained with historical data from a selected public building. The dataset includes sensor measurements and satellite solar radiation inputs, which allows models to capture both internal and external influences while measuring indoor temperature changes. The PINN model is mainly designed to integrate a physics-based residual term and find the physics loss by comparing with the predicted data. So, this ensures the model maintains physical plausibility even if there is noisy data in the scenarios.

## 3.2 Data Preprocessing

In this research, before training any ML or physics-informed models, one of the main critical steps is to preprocess the raw data and ensure usability, consistency, and quality. In this research, indoor temperature prediction relies on multiple data sources, including sensor-based data from buildings and external environmental data such as satellite based solar radiation parameters. These datasets vary in terms of sampling frequency. So, some preprocessing operations were required to process the data.

### **Handling Missing Data:**

It is known that in the real world, sensor data often has missing and corrupted values due to some transmission errors, sensor faults, or downtime. The dataset used for this research also has some time intervals with missing values, and there was no continuation of values, so this cause time alignment issues, which are addressed by applying spline interpolation at a fixed 5-minute interval. This spline interpolation approach generates smooth curves with existing data points and also estimates missing values with better accuracy than other linear methods. So, this interpolated dataset mainly ensures uniform time intervals and fills gaps in the data sequence.

### **Timestamp Alignment:**

The telemetry data from sensors sometimes included different time zones and formats. As for this research, satellite data is also used, so for all timestamps, it was converted to a uniform format, though sensors were established in Pietarsaari, Jakobstad, Finland, and we have downloaded satellite data from other sources, so it was localized to the desired time zone UTC+2 for Finland. So, this step mainly ensures consistency when merging datasets from different sources.

**Feature selection:**

The dataset includes many variables, but for this research, it was found that only a subset was relevant for temperature prediction modeling. So, the selected features included:

- Indoor temperature (Target variable)
- Supply Air Temperature
- Outdoor temperature
- Satellite-based solar radiation values

**Data Normalization:**

For bringing all features to the same scale and avoiding any dominance by any single variable, the Robust scaler from scikit learn was used. Without normalization variables like radiation, outdoor temperature, and supply air, return air temperature would appear on different numerical scales. This can cause models, mainly neural network models, to give more weight to features with large values simply because of their scale. For ML tasks, two popular scaling methods are: MinMax scaler and Standard scaler.

**MinMax Scaler:** It mainly scales values to a fixed range, typically [0,1]. This scaler method is simple, and it works well when the data is clean. But it is very sensitive to any outliers. A single extreme value can stretch the range, and there is a chance to distort all other values.

**Standard Scaler:** It mainly transforms data using the mean and standard deviation. So it centers the data around zero with unit variance. However, it also sometimes struggles when outliers are present because the mean and standard deviation are easily influenced by extreme values. It is already known that in environmental and building telemetry datasets, outliers are almost unavoidable. So sudden sensor spikes, sensor drops can happen, and it is normal system behavior.

Because of this, we mainly chose the Robust Scaler method. This method is designed to reduce the effect of outliers. Instead of using the mean and standard deviation, it uses the median and the interquartile range (IQR), which is more stable when unusual values appear.

The scaling formula is:

$$X_{scaled} = \frac{X - median(X)}{IQR(X)} \quad (1)$$

So, by using this way, even if a sensor suddenly reports an unrealistic peak, the scaling process stays stable and meaningful.

Overall, Robust Scaler ensured that all the features were on a similar scale without any sudden anomalies or distortions in the learning process, which leads to more reliable gradient updates and better model convergence.

### 3.3 Dataset Characteristics and Limitations

The dataset used in this research mainly a collection of indoor temperature, outdoor temperature, supply air temperature, and satellite-based radiation values which is collected from a public building of Pietarsaari, Jakobstad, Finland. This dataset mainly covers the time between November 12, 2024, and January 7, 2025. Although, this dataset is important for developing and testing indoor temperature prediction models, it is also very important to understand its limitations to the proper assessment of model performance.

At first It is important to note that the exact placement of sensors inside the building is unknown. Indoor temperature sensors can be located near windows, doors, and ventilation outlets. This kind of situation can influence in the recorded of temperature readings. As there is an absence of detailed sensor documentation, it is tough to assess the accuracy of the measurements across the entire indoor environment.

Second, while data exploration several sudden spikes and abnormal jumps were observed in the indoor temperature readings. These types of sudden spikes which are mainly anomalies may be caused by multiple factors, such as:

- Errors in sensor calibration
- Temporary heating or cooling system adjustments issue,
- Window or door opening events,
- Data transmission related issue,
- Rare environmental disturbances.

Since indoor temperature normally changes slowly over short period of time, but if there is any sudden changes that may indicate sensor noise or operational errors. Although interpolation and smoothing techniques were applied but still such anomalies may affect model accuracy.

Furthermore, the dataset has only one indoor temperature sensor, which makes it difficult to perform any cross-validation between multiple sensors. Having several indoor sensors at different locations would allow us to identify if there is any faulty readings, spatial averaging, or multizone modeling. The single sensor makes fault detection challenging and it increases dependence on preprocessing methods to ensure data quality.

Finally, the short duration of the dataset, however, limits the models exposure to longer seasonal variations like extreme cold conditions, or unusual summer conditions, and different operational states of the HVAC system. Although the dataset is sufficient for testing the feasibility of short-term forecasting and physics-informed learning, the long-term thermal behavior of the building may not be fully captured by it.

### 3.4 Feature Engineering

Feature engineering is a process of selecting and creating useful input variables that help ML models to produce better predictions. In this research, it plays an important role in improving the accuracy of indoor temperature forecasting models.

From the original telemetry building dataset from Pietarsaari, Jakobstad, Finland. We have selected some key variables such as indoor temperature ( $T_{in}$ ), Outdoor temperature ( $T_{out}$ ), Supply air temperature, and satellite-based solar radiation. It is found that these are the important factors that normally affect the temperature of a building. However, it is important to clarify that these features were selected from the available sensor data and not necessarily based on a full ranking of importance. Therefore, it is more correct to say that these variables are assumed to be some of the most important.

Additionally, the dataset for this building was missing two critical variables information such as central heating water temperature and district heating flow and both of which are critical variables in understanding heat delivery and energy consumption. So, without these variables it is difficult to directly calculate heating energy or differentiate between internal and external heating influences.

To make the model easy to understand and the heating and cooling effect better, we created a new feature called ventilation temperature difference ( $\Delta T$ ). It is calculated as  $T_{supply} - T_{indoor}$ . This  $\Delta T$  refers to the temperature difference between the supply air temperature and the indoor air temperature. There is also another temperature difference, which is important difference between the indoor temperature and the outdoor temperature. Another important feature is that we have used step by step recursive forecasting and it uses the last timestep's data (temperature at time  $t$ ) to predict temperature at time  $t+1$  and then it uses the predicted value at  $t+1$  as part of the input to forecast  $t+2$  kind of one-step-ahead predictions.

Before training the models, we scaled all the features using RobustScaler from preprocessing library of SKLearn toolbox . This mainly helps to normalize the data and makes the training more stable, especially if there are any unusual spikes or outliers in the dataset. We chose Robust Scaler specifically because it is less influenced by outliers compared to other scalers like MinMax Scaler or Standard Scaler. So instead of using mean and standard deviation, Robust scaler uses median and interquartile range. It is found that it is more reliable when the dataset contains some extreme values. In our research case, it is sudden temperature change or sensor spikes, which are common in building telemetry data.

<b>Variable</b>	<b>Mean</b>	<b>Median</b>	<b>Std. Dev</b>	<b>Min</b>	<b>Max</b>
<b>Indoor Temperature (°C)</b>	22.15	22.27	2.30	13.26	28.02
<b>Supply Temperature (°C)</b>	34.65	33.65	7.31	20.48	123.36
<b>Outdoor Temperature (°C)</b>	-1.78	-1.45	4.55	-14.99	9.23
<b>Satellite Radiation</b>	4.40	0.00	11.24	0	108.91

**Table 1 : Summary Statistics of Input Variables and Justifying the Use of RobustScaler**

So here the table includes the mean, median, standard deviation, minimum, and maximum values for each feature. and satellite radiation and supply air temperature variables show a large difference between their minimum and maximum values with standard deviations of 11.24 and 7.31 respectively. it is a indication presence of significant outliers which can negatively impact model performance if not handled properly. Similarly for the indoor temperature the range is from 13.26 to 28.02, which as we discussed earlier about sudden spikes and it supposed to be more stable within short timeframes. So, this observation justify that instead of using StandardScaler or MinMaxScaler, we applied RobustScaler which uses the median and interquartile range and is less sensitive to outliers.

Overall, this feature engineering process helps us to prepare a more meaningful dataset which mainly supports both ML and PINN models for predicting indoor temperature more accurately.

### 3.4.1 Selection of relevant features

Based on the literature review and domain knowledge, a few variables were selected as we discussed in the previous section. In this section, we will discuss a little about those features as they were used for indoor temperature prediction:

**Supply Air temperature:** It shows the amount of heat air entering the space, and it has a direct impact on the indoor temperature.

**Outdoor temperature:** It captures the external weather impact on indoor thermal conditions, especially as the research is done in Nordic climates.

**Solar radiation:** It is used as a proxy for solar gain. High solar radiation typically lead to increased indoor temperatures, particularly in buildings with large windows or poor insulation.

Although central heating water temperature data would have been an important feature for modeling the indoor temperature mainly for estimating heat transfer through radiators. But unfortunately, it was not available for each room or zone in our dataset. So, it could not be included in this research.

### 3.4.2 Target variable creation

In this research, the outcome variable is indoor temperature. Therefore, the model is trained to predict future temperatures. Instead of limiting the model for a single step or short-term prediction, the models were designed to estimate indoor temperature at, at 1-hour, 6-hours, and 12-hours intervals ahead.

This research approach mainly allows for both short-term control optimization and medium-term planning. These are essential in case of HVAC operations effectiveness and

energy management. This forecasting research can be treated as a regression problem, where the model results are a single continuous temperature value corresponding to a specific prediction horizon.

### **3.4.3 Temporal Structure for Forecasting**

In this research, several experiments were conducted, and in the final version of the experiments, the forecasting approach was simplified to use only the most recent observation (last step) instead of a 30-minute or 60-minute loopback window. This decision was made based on performance evaluation and model interpretability.

Each input to the model consists of the latest available values of indoor temperature at time  $t$ . For predicting the indoor temperature at future time steps  $t+60$ ,  $t+360$ , or  $t+720$  minutes. This one-step input structure reduces computational complexity while still retaining meaningful prediction. Moreover, it aligned with operational use cases where real-time decisions are made based on the last sensor readings. Despite the absence of any temporal sequence in the modelling, this formulation sequence allows for effective learning for any immediate impact on indoor thermal conditions.

Long term sensor history was not assumed to be very meaningful in this research as in this context, that since outside weather and indoor conditions are often non-stationary. So, relying on long sequences could lead to overfitting patterns and therefore it can be said that time series prediction can merely utilize some regular usage patterns, but not otherwise be very useful.

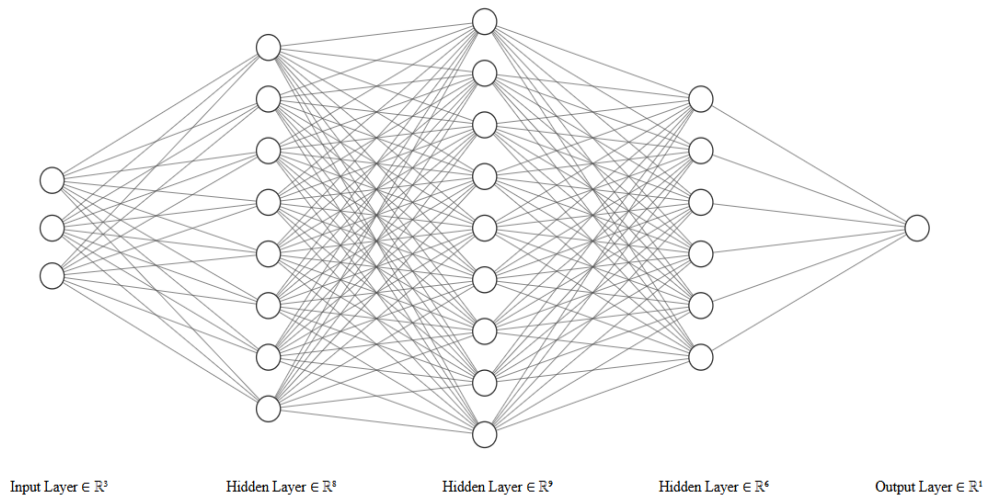
## **3.5 Model Architecture**

In this research to evaluate the performance of different forecasting strategies for indoor temperature prediction, several ML models were developed and trained in this research experiment. The selected models mainly represent a combination of traditional regression techniques, neural networks, and PINNs. All models were trained using the

same preprocessed dataset and temporal sequence to ensure comparability between all the models.

### 3.5.1 Feedforward Neural Network

The FNN model was mainly implemented using PyTorch, where multiple fully connected layers were explored with ReLU activations. Here three key variables: supply air temperature, outdoor temperature, and satellite radiation value were taken as a input. And in the model, an exhaustive search over single, double, and triple-layer configurations was performed. In each layer, a search was performed varying the number of neurons from 1 to 10 per layer. Leading to almost 1110 architecture combinations being checked to identify the most optimal architecture of neural networks which is layers of 8, 9, and 6 neurons respectively. Then the final output layer contains a single neuron which is the predicted indoor temperature value for the next time step. Each model was trained to minimize the MSE between the predicted and actual indoor temperature values. Finally, the best performing architecture was selected based on validation performance using early stopping with a certain patience value.



**Figure 1 Architecture of the Feedforward Neural Network used for Indoor Temperature Prediction.**

### **3.5.2 Long Short-Term Memory**

LSTM networks are a specialized neural network architecture designed for time-series prediction. They can retain information from previous time steps through internal memory cells. In this research, The model architecture used in this study consisted of an input layer, two stacked LSTM layers (with 64 and 32 units), followed by two dense layers for final regression output. Unlike FNN models, LSTM models need sequences of multiple time steps. In this research, each training example was constructed from 12 previous time steps of data to predict the temperature at the next five-minute interval. The input features included supply air temperature, outdoor temperature, satellite-derived radiation, and indoor temperature. For data splitting, the LSTM model followed the same timestamp-based separation used for all other models. The final 288 samples (one day) was reserved as the test set, while all other samples used for the training set. The input sequences were scaled using RobustScaler, which is consistent with the preprocessing that was applied to the other ML models. The model was trained using the Adam optimizer and a mean squared error (MSE) loss function. To prevent overfitting, early stopping was applied, and the network was trained for up to ten epochs with a batch size of 32.

### **3.5.3 Extreme Gradient Boosting**

XGBoost model is a robust ensemble learning method, which is based on gradient boosted decision trees, and it was used for its effectiveness in handling structured tabular data. In this research, the model was mainly trained using `reg:squarederror` as the objective function, which corresponds to the standard squared error loss used for continuous regression tasks. The reason for using `reg:squarederror` is that indoor temperature forecasting is a continuous regression problem. The square error loss penalizes larger mistakes more strongly than smaller ones. L2 and L1 regularization work as a penalty to reduce the impact of overly large leaf weights in the boosted trees. This loss function penalizes larger prediction errors more strongly, making it ideal for indoor

temperature forecasting, where minimal deviation from the actual temperature is crucial. Furthermore, the model incorporates L1 and L2 regularization to prevent model overfitting by penalizing overly complex tree structures.

The model uses the most recent available sensor readings to make predictions. These readings include supply air temperature, outdoor temperature, satellite solar radiation, and the previous indoor temperature value ( $t-1$ ). These features together represent the thermal factors that influence indoor temperature behavior. XGBoost then learns nonlinear relationships based on these inputs and predicts the next indoor temperature value. This forms the basis for forecasting at 1-hour, 6-hours, and 12-hours horizons in this thesis.

#### **3.5.4 Random Forest**

The RF regressor modeling includes a classical ensemble baseline. So it was trained with 500 decision trees, a maximum depth of 15, and bootstrap sampling, with node splits controlled by  $\text{min samples split} = 5$ ,  $\text{min samples leaf} = 2$ . In the implementation of modelling, the same data split was done as it was implemented for other models. This approach provided a useful and meaningful comparison against more complex neural and hybrid models. The hyperparameters for the RF model were chosen through a combination of best-practice defaults, and the characteristics of the dataset.

All these models were evaluated under a multiple-horizon forecasting setup, so each model predicts the indoor temperature at 1-hour, 6-hours, and 12-hours intervals using the latest available indoor temperature and other features as an input.

### **3.6 Physics-Informed Neural Network Modeling**

In conventional ML models, it can capture patterns from data. It often lacks physical interpretability, and there is a chance that it may produce unrealistic predictions when extrapolated beyond training conditions. For this limitation, A PINN approach was

integrated into the research. This method combines data-driven learning with known physics laws, mainly a simplified equation for indoor temperature dynamics. The Fourier law of heat transfer explains how an individual object's temperature changes over time when exposed to an environment with a different temperature.

Fourier law of heat transfer definition: The rate of change of temperature of an objects temperature is directly proportion to the difference between its temperature and the surrounding ambient temperature.

Mathematical Form:

$$\frac{dT}{dt} = -k(T - T_{ambient}) \quad (2)$$

Where:

- T: Temperature of the object (indoor temperature)
- $T_{ambient}$ : Surrounding or environmental temperature (outdoor temperature)
- k: A positive constant depending on the system (thermal conductance)
- $\frac{dT}{dt}$ : Rate of temperature change

### 3.6.1 Governing Differential Equation

This research adapts the governing equation used in (Hannula et al., 2025; Häkkinen, n.d.). In our research, the equation is reinterpreted to model air-based systems by replacing the supply water temperature with supply air temperature. The rest of the structure remains consistent, maintaining the balance between internal heat gains, losses, and solar radiation.

The equation is formulated as follows:

$$\frac{dT}{dt} = \theta_0 \cdot (T_{supply\ air} - T_{in}) + \theta_1 \cdot (T_{in} - T_{out}) + \theta_2 \cdot \Phi \quad (3)$$

Where:

- $T_{in}$ : indoor temperature

- $T_{\text{supply\_air}}$ : supply air temperature
- $T_{\text{out}}$ : outdoor temperature
- $\Phi$ : solar radiation
- $\theta_0, \theta_1, \theta_2$ : learnable parameters representing heat transfer rates from respective sources

The term  $\theta_1(T_{in} - T_{out})$  directly represents Fourier Law of Cooling, which represent that indoor temperature will decrease if it is warmer than outdoors, and the rate depends on  $\theta_1$ . The ODE and the approach was adopted in (Hannula et al., 2025; Häkkinen, n.d.). Which successfully integrated physical laws into learning indoor temperature patterns.

### 3.6.2 Learning Physics Parameters

In the equation, there are three parameters  $\theta_0, \theta_1, \theta_2$  and all of these were learned directly from the data using gradient descent. Here, the time derivative  $\frac{dT_{in}}{dt}$  was approximated using first-order finite difference (Euler's Method), which is calculated as:

$$\frac{T_{in}(t + 1) - T_{in}(t)}{\Delta t} \quad (4)$$

Where:

- $T_{in}$  = Indoor temperature at a specific moment in time
- $T_{in}(t + 1)$  = Indoor temperature at the next step.
- $T_{in}(t)$  = Indoor temperature at the current time step
- $\Delta t$  = The time difference between two consecutive measurements.

A loss function which was then defined to minimize the mean squared error between the left-hand side of the derivative and the right hand side of the equation. So, after a

few iterations, these three learnable parameters provided a physically meaningful simulation of indoor temperature evolution.

### 3.6.3 Physics Loss Integration

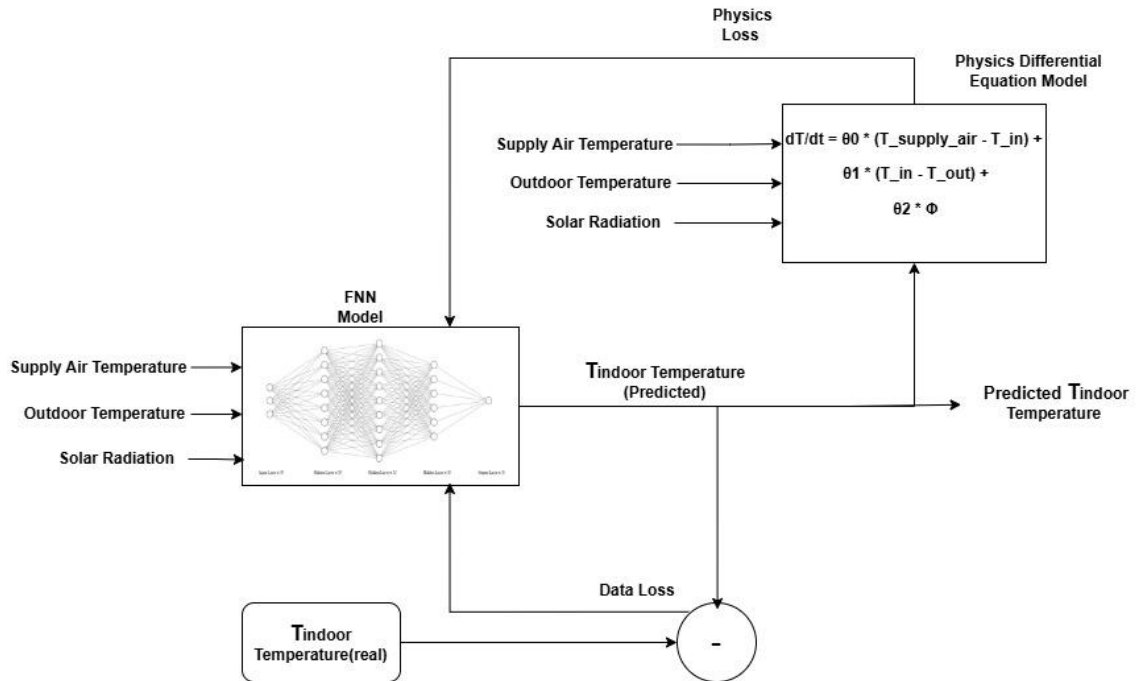
For incorporating the physical knowledge into the neural network model a composite loss function was defined:

$$\mathcal{L}_{total} = \lambda_{data} \cdot \mathcal{L}_{data} + \lambda_{physics} \cdot \mathcal{L}_{physics} \quad (5)$$

Where:

- $\mathcal{L}_{data}$ : Standard MSE loss between predicted and sensor indoor temperature
- $\mathcal{L}_{physics}$ : MSE between model dynamics and the differential equation
- $\lambda_{data}, \lambda_{physics}$ : Weighting factors balancing the contribution of data and physics terms.

In this research,  $\lambda_{data} = 1$ , and  $\lambda_{physics} = 0.003$  were selected while code implementation based on experiment which will maintain an appropriate balance between data fitting and physical realism.



**Figure 2: Architecture of the Physics-Informed Neural Network (PINN) for indoor temperature prediction**

### 3.6.4 Machine Learning Model Comparison

In this research, both traditional and data driven ML model was developed and using learnable theta parameters and the Euler method a forward simulation also conducted for prediction temperature evolution. This simulated result was compared with the actual measurements and other ML models like FNN, LSTM, XGBoost. So, it can evaluate how well the physics model aligned with other observed dynamics.

So, the PINN model not only provide almost competitive accuracy but also it ensures that the predictions obeyed fundamental thermodynamic principles. This total process mainly enhanced the model robustness particularly in some cases where data was sparse.

### 3.7 Evaluation Metrics and Performance Comparison

For assessing the predictive accuracy and generalization capabilities of developed ML models, a commonly used regression evaluation metric was employed. These metrics

mainly provided the insights into both absolute error and overall goodness of fit between the predicted and actual indoor temperature values.

### 3.7.1 Mean Absolute Error

The MAE is a statistical metric which is mainly used to measure the average magnitude of the errors between predicted and observed values. It is defined by:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_{true} - y_{pred}| \quad (6)$$

Here,  $y_{true}$  is actual indoor temperature and  $y_{pred}$  is predicted temperature, and  $n$  is the number of samples. MAE gives an intuitive measure of model accuracy in degree Celsius. This is the reason that it is suitable for thermal comfort temperature evaluation.

### 3.7.2 Root Mean Squared Error

The RMSE mainly penalizes larger error more than MAE due to the squaring operation, as it offers more sensitive evaluation metric when larger deviation are more critical. It is defined as:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_{true} - y_{pred})^2} \quad (7)$$

RMSE is particularly useful as when there is large deviation from true values are less tolerable. Here,  $y_{true}$  is the actual indoor temperature and  $y_{pred}$  is the predicted temperature.

### 3.7.3 Coefficient of Determination

The coefficient of determination ( $R^2$ ) score evaluates how well the predictions approximate with the actual values. It mainly represent the proportion of variance in target variable which is predictable from the input features. It is defined as:

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_{true} - \bar{y}_{pred})^2}{\sum_{i=1}^n (y_{true} - \bar{y}_{true})^2} \quad (8)$$

Here,  $\bar{y}$  is the mean of actual target values. So,  $R^2$  score is close to 1 indicates a strong agreement between the prediction and the ground truth.

## 4 Results and Findings

This chapter presents the experimental results. These results are from the implementation and evaluation of three different modeling strategies. These strategies are for indoor temperature prediction. FNN, ODE model based on physics, and PINN. All models were trained and tested using a dataset containing 16,089 samples collected at five-minute intervals between November 12, 2024, and January 7, 2025. Of these samples, 15,801 were used for training and 288 for testing. All models used the same input variables: supply air temperature, outdoor temperature, and solar radiation, and here indoor air temperature was the target variable. The forecasting task was performed for 1-hour, 6-hours, and 12-hours horizons.

### 4.1 FNN Architecture Search and Performance

An large scale architecture search was conducted across 1,110 different FNN architectures. The search varied the number of hidden layers (from 1 to 3) and the number of neurons per layer (from 1 to 10) while using the rectified linear unit (ReLU) activation function. The goal was to identify the optimal configuration based on the validation performance metrics of MAE, RMSE, and  $R^2$  score.

Parameter	Value
Hidden layers	1 to 3 layers
Neurons per layer	1 to 10
Total architectures tested	1,110
Activation function	ReLU
Optimizer	Adam
Loss function	MAE
Scaler	RobustScaler

**Table 2: FNN Architecture Search Parameters Used in This Study**

top 20 architectures by MAE results :

Architecture	Type	Parameters	MAE	RMSE	R <sup>2</sup>
[8, 9, 6]	3-layer	188	0.383845	0.635125	0.617929
[7, 7, 1]	3-layer	101	0.396185	0.590809	0.669387
[9, 6, 4]	3-layer	138	0.402752	0.709243	0.523551
[10, 2, 10]	3-layer	113	0.423718	0.653618	0.595355
[10, 6, 9]	3-layer	189	0.429039	0.667377	0.578140
[7, 8, 3]	3-layer	130	0.431974	0.656843	0.591352
[7, 7, 7]	3-layer	155	0.438993	0.629608	0.624537
[8, 1, 10]	3-layer	80	0.458709	0.667606	0.577850
[10, 3, 8]	3-layer	124	0.496059	0.743952	0.475776
[8, 8, 5]	3-layer	163	0.498181	0.675320	0.568038
[8, 5, 9]	3-layer	149	0.499871	0.680575	0.561289
[8, 9, 7]	3-layer	199	0.504241	0.703032	0.531859
[8, 10, 9]	3-layer	239	0.507396	0.759271	0.453965
[8, 8, 10]	3-layer	213	0.512494	0.756914	0.457350
[10, 7, 9]	3-layer	209	0.529363	0.748422	0.469458
[6, 5, 9]	3-layer	129	0.532644	0.723839	0.503739
[7, 6, 4]	3-layer	116	0.536624	0.764388	0.446581
[6, 7, 8]	3-layer	152	0.538013	0.777529	0.427389
[6, 9, 7]	3-layer	171	0.538466	0.877027	0.271462
[10, 7, 10]	3-layer	218	0.543124	0.718554	0.510959

**Table 3 :Top 20 FNN Architectures Based on Validation MAE**

From this search, we identified the best-performing architecture as [8, 9, 6], 3-layer FNN with 188 trainable parameters. Its validation performance was MAE: 0.3838 °C, RMSE: 0.6351 °C, R<sup>2</sup>: 0.6179. After being retrained on the full training dataset, the model achieved the following performance on the full test set test MAE: 0.7255 °C, test RMSE: 0.9257 °C, test R<sup>2</sup>: 0.1884. While the FNN produced reasonable single-step predictions, its recursive multi-step forecasting stability still needed to be looked into more.

## 4.2 One-Hour Prediction Window

The best-performing FNN architecture was evaluated using a 1-hour one-step-ahead predictions corresponding forecasting window, which corresponded to 12 sequential five-minute prediction steps. In this setting, the model's predictions are used as inputs at each step, allowing error accumulation to be observed.

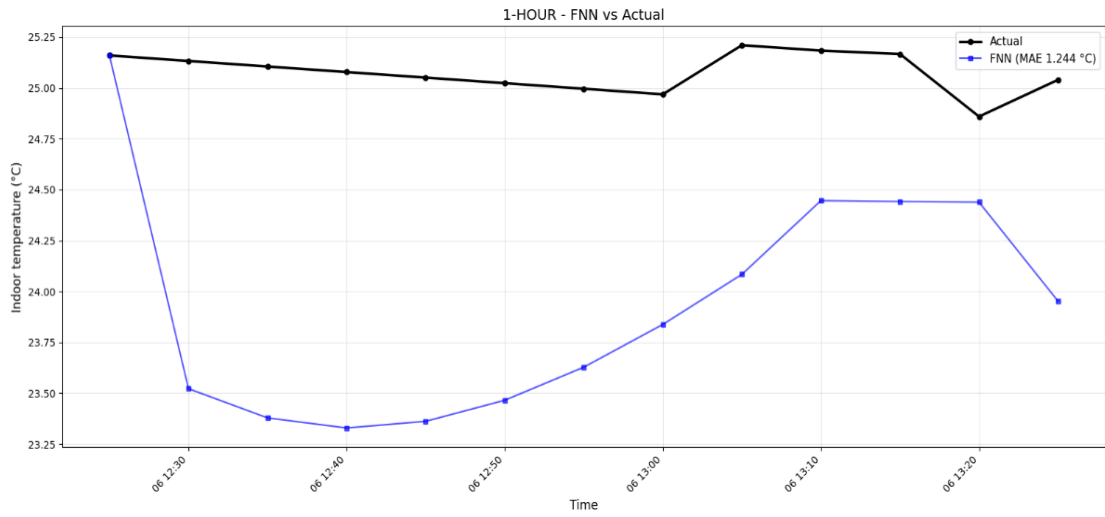


Figure 3 : One Hour Prediction: FNN vs Actual

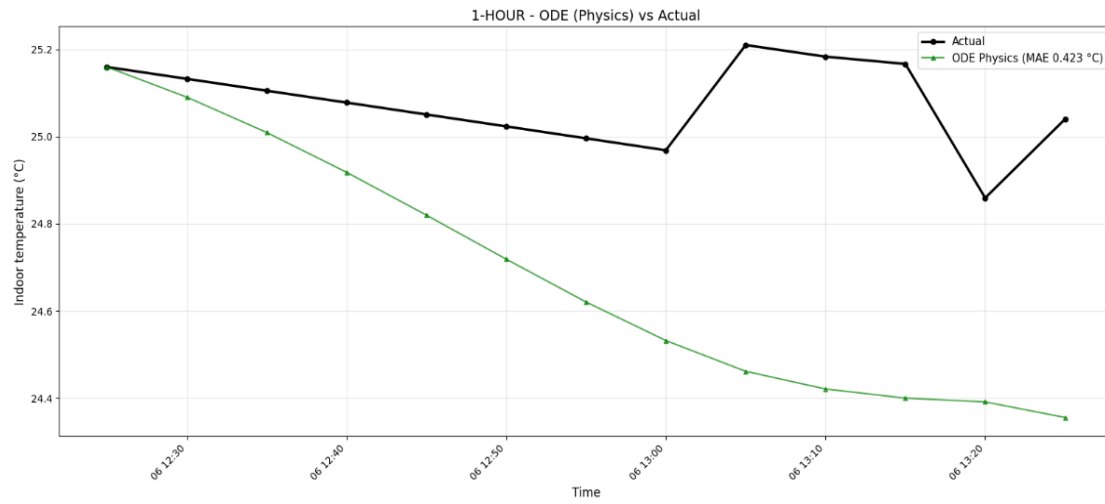


Figure 4: One Hour Prediction: ODE(Physics) vs Actual

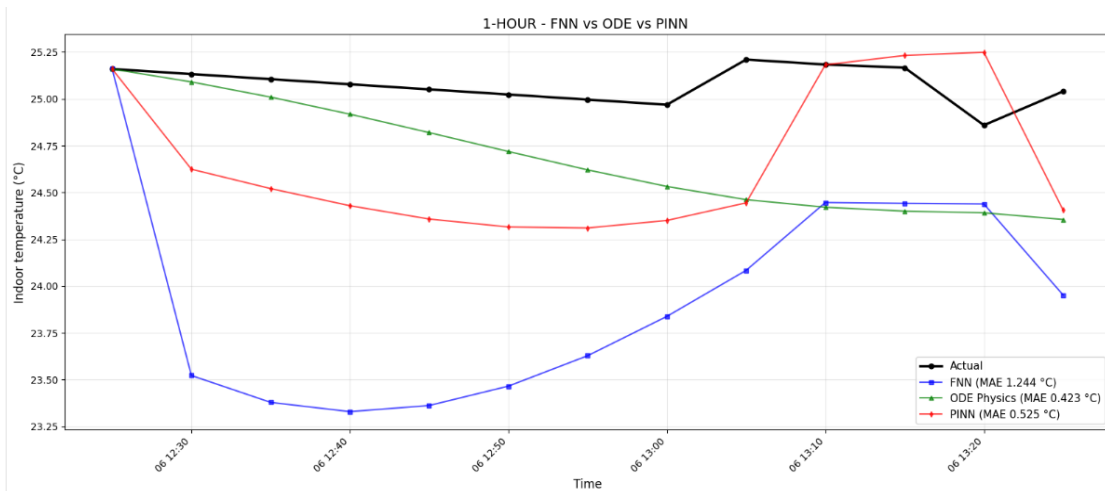


Figure 5: One Hour Prediction: FNN vs ODE vs PINN

Model	MAE (°C)	RMSE (°C)	R <sup>2</sup>
FNN	1.24	1.31	-1
PINN	0.52	0.57	-1

**Table 4: Performance Metrics for the One Hour Prediction (FNN vs PINN)**

Here, the results clearly show the limitations of a purely data-driven model when used for recursive multi-step prediction.

The FNN shows a rapid increase in errors over the short 1-hour period. This results in a high MAE and a negative R<sup>2</sup> value, which indicates that the model's performance is worse than a simple mean baseline. This behaviour is expected because recursive forecasting increases small deviations in single-step predictions, specifically when the model lacks explicit knowledge of the underlying thermal dynamics. However, the physics-based ODE model achieves the best short-term accuracy, since the ODE model is based on the physical structure of heat transfer, it remains stable over recursive predictions and avoids the drift observed in the FNN. But in the case of PINNs, it performs substantially better than the FNN and approaches the accuracy of the ODE. Its ability to combine data-driven learning with the equation prevents instability and maintains stable forecasting behaviour. Although the PINN is slightly less accurate than the pure ODE model in the short term, it has the advantage of learning from data while maintaining physical constraints, which is beneficial in the long term.

### 4.3 Six-Hours Prediction Window

To evaluate model stability over a longer forecasting horizon, the FNN, ODE, and PINN models were tested over a 6-hour period.

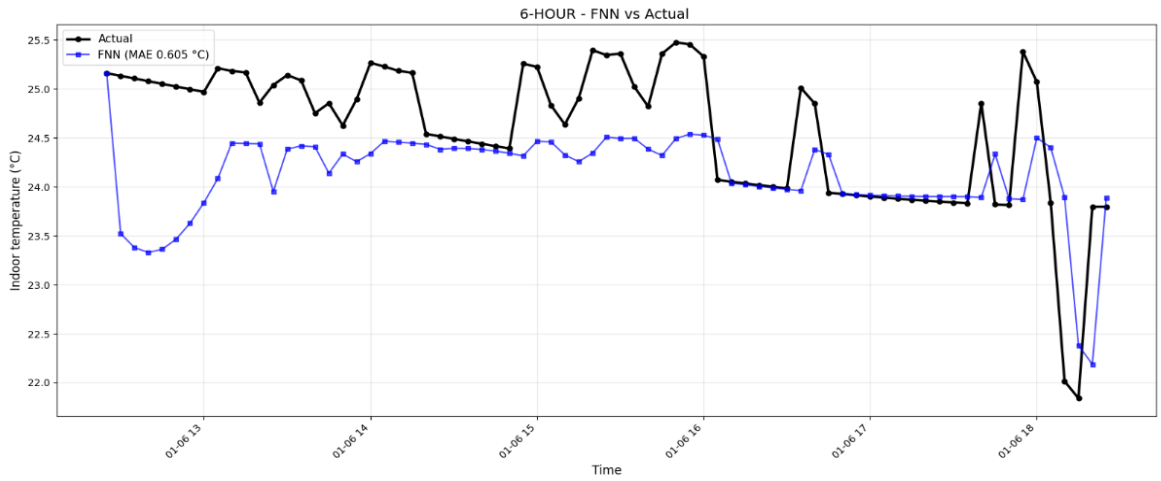


Figure 6: Six-Hours Forecast: FNN vs Actual

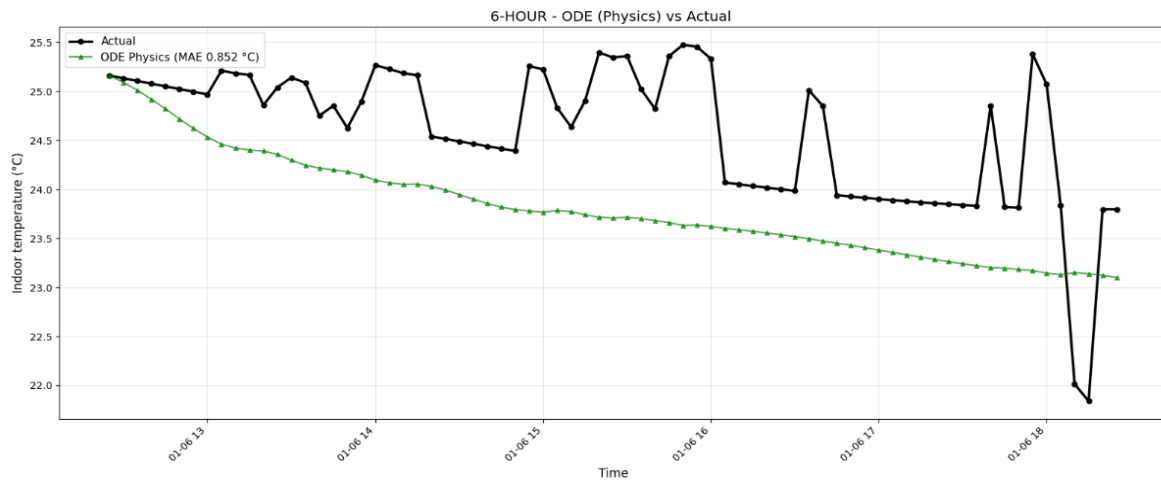


Figure 7: Six-Hours Forecast: ODE vs Actual

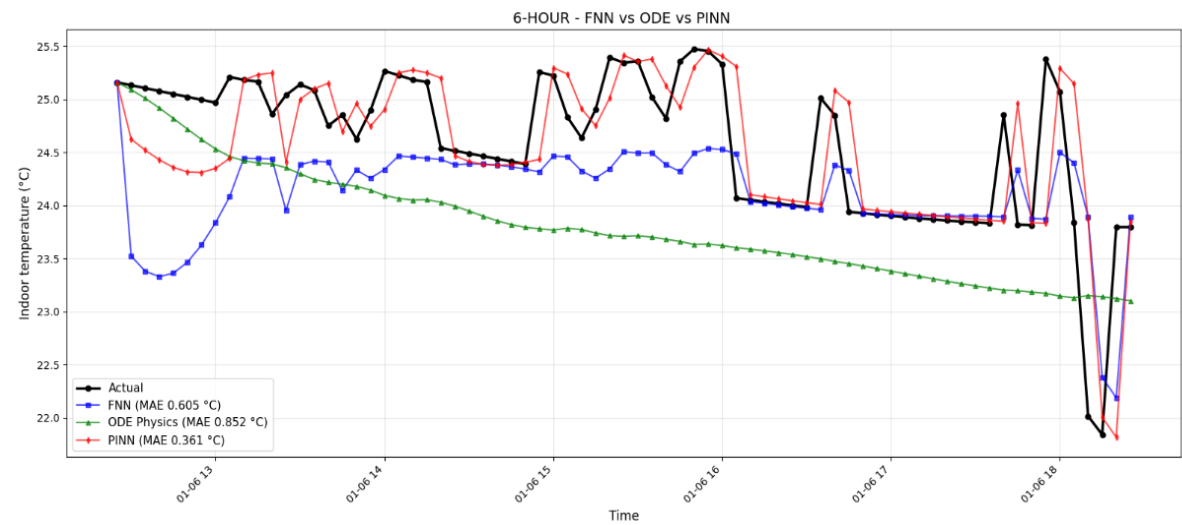


Figure 8: Six-Hours Forecast: FNN vs ODE vs PINN

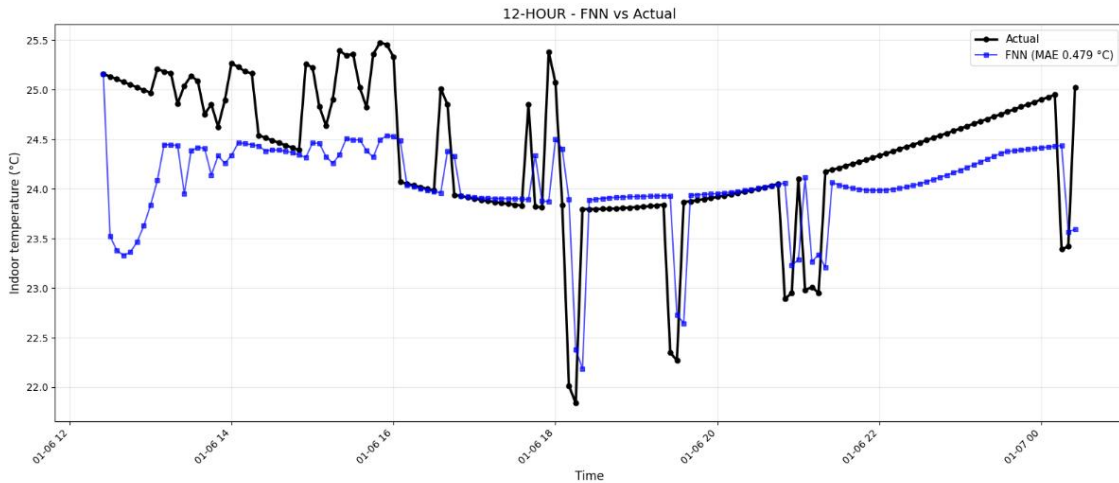
Model	MAE (°C)	RMSE (°C)	R <sup>2</sup>
FNN	0.60	0.62	-1
PINN	0.36	0.34	0.34

**Table 5: Performance Metrics for the Six Hours Prediction (FNN Vs PINN)**

Here, the results show a clear shift in model performance compared to the one-hours window. So according to these results, the FNN performs slightly better with reduced MAE and root mean squared error. But the negative R<sup>2</sup> value shows that the model still fails to generalize the underlying physical behaviour. For ODE model, the performance of the model decreases over longer horizons. Here, PINNs achieves the best performance on the 6-hour window with positive R<sup>2</sup>, indicating meaningful predictability.

#### 4.4 Twelve-Hours Prediction Window

The final evaluation scenario examines the models behavior over a 12-hours period.



**Figure 9: Twelve-Hours Prediction: FNN vs Actual**

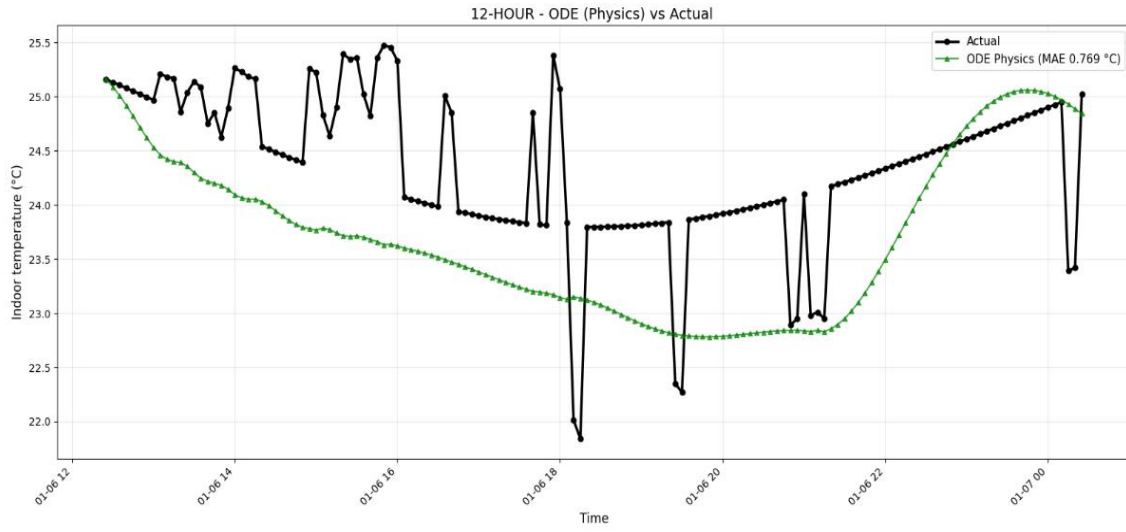


Figure 10: Twelve-Hours Forecast: ODE vs Actual

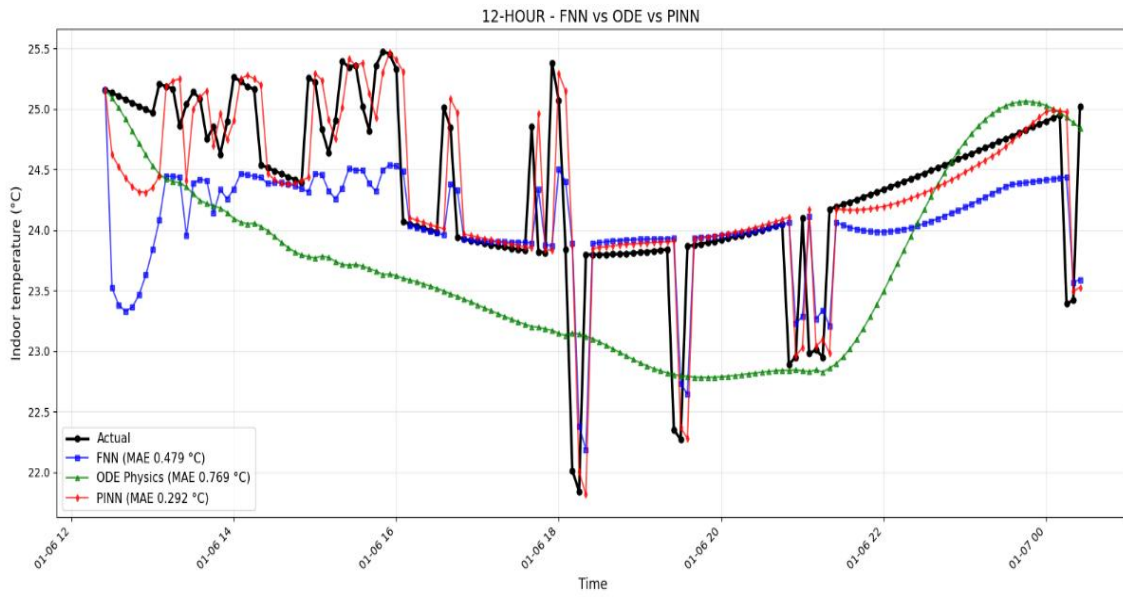


Figure 11: Twelve-Hours Forecast: FNN vs ODE vs PINN

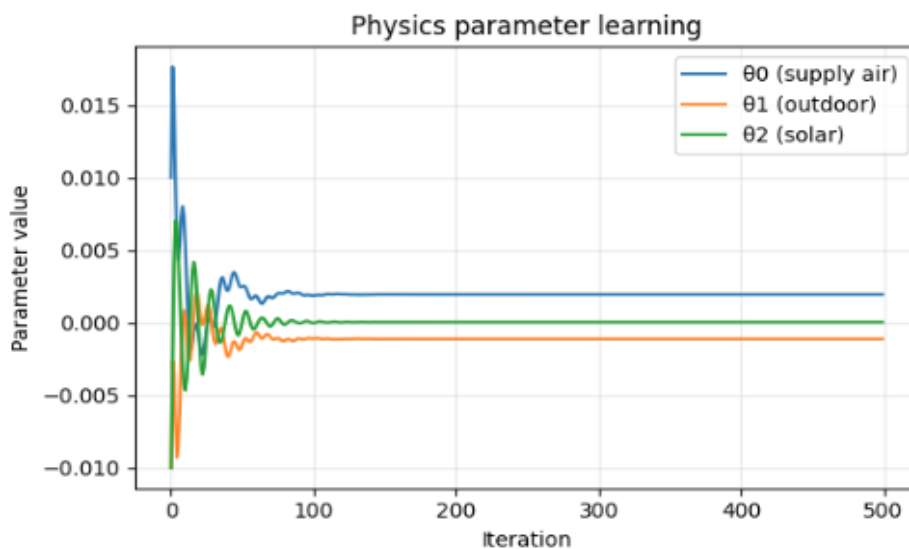
Model	MAE (°C)	RMSE (°C)	R <sup>2</sup>
FNN	0.47	0.68	-0.29
PINN	0.29	0.51	0.63

**Table 6: Performance Metrics for the Twelve Hours Prediction (FNN Vs PINN)**

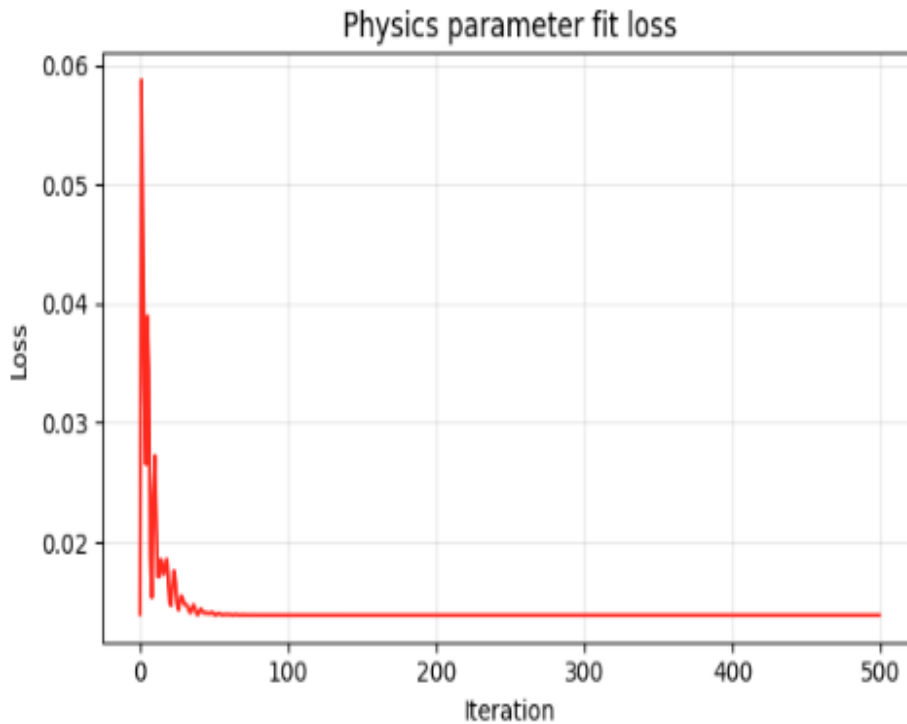
In this, 12-hours analysis clearly shows differences in model stability. The FNN model continues to accumulate recursive errors, resulting in divergence and negative R<sup>2</sup> values. The ODE model remains stable and PINNs gives the most accurate and consistent long-term predictions. It has the lowest MAE and RMSE, and it is the only one with a strongly positive R<sup>2</sup>. These results confirm that integrating physics conditions results in far superior long-term prediction performance compared to purely data-driven or physics-based models.

#### 4.5 Physics-Based ODE Model Performance

The physics-based ODE model is easy to understand because it is based on Fourier law of heat conduction. It uses a simple equation that shows how indoor temperature changes with supply air temperature, outdoor temperature, and solar radiation.



**Figure 12: Physics Parameter Learning**



**Figure 13: Physics Parameter Loss Over Epochs**

Parameter	Value
$(\theta_0)$ (supply air effect)	0.0019
$(\theta_1)$ (outdoor influence)	-0.0011
$(\theta_2)$ (solar influence)	0.00005

**Table 7: Physics Learnable Parameters Value**

These values match the expected thermal behaviour. As we know Supply air has a warming effect, the outdoor temperature contributes to cooling, and solar radiation provides a small, positive heat gain. In ODE model, it produces smooth and physically meaningful predictions and maintains stability during long-horizon prediction. However, in the models it cannot fully capture rapid variations and there is reason that in the dataset we have sudden outliers which is not possible for physics-based equation to

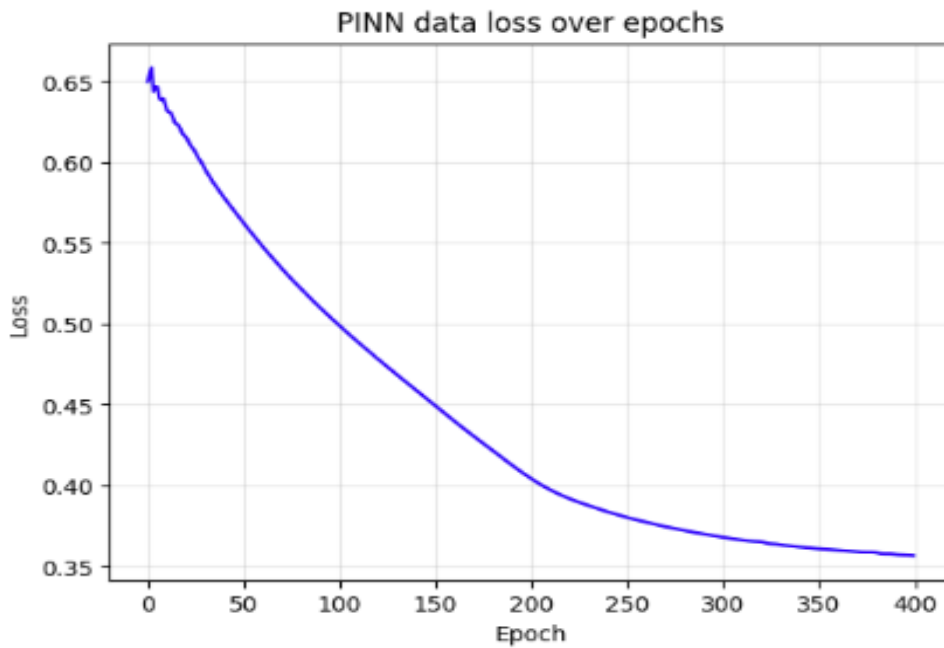
identify as our input features does not contain that relevant information. Still, it provides a reliable physical baseline and is an essential component for the PINN framework.

#### 4.6 Physics-Informed Neural Network Performance

The PINN combines learned differential equation-based physics with a FNN architecture. This hybrid modeling approach allows the PINN to learn complex nonlinear patterns from the dataset while maintaining physical consistency.

Parameter	Value
Input variables	Tsup, Tout, $\Phi$ , Tin
Physics equation	$dT/dt = \theta_0(T_{sup}-T_{in}) + \theta_1(T_{in}-T_{out}) + \theta_2\Phi$
Learnable parameters	$\theta_0, \theta_1, \theta_2, C$
Network architecture	[8, 9, 6]
Loss function	Data loss + Physics loss
Optimizer	Adam

**Table 8: PINN Architecture Parameter Used In This Study**



**Figure 14: PINN Data Loss Over Epochs**

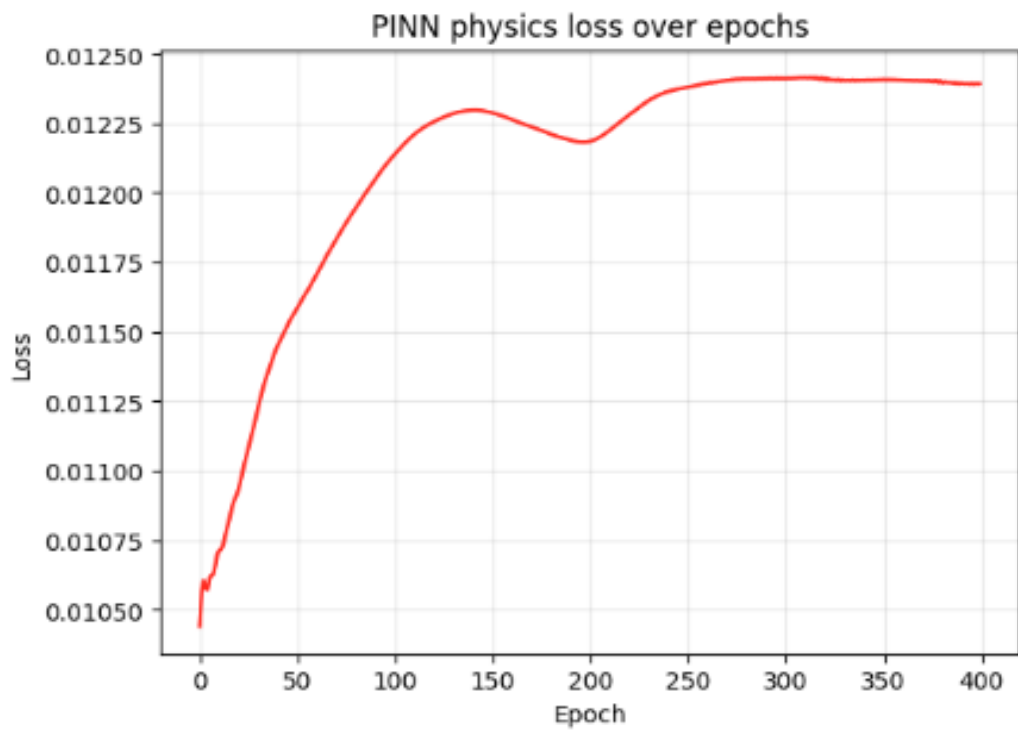


Figure 15: PINN Physics Loss Over Epochs  
Data + Physics Loss

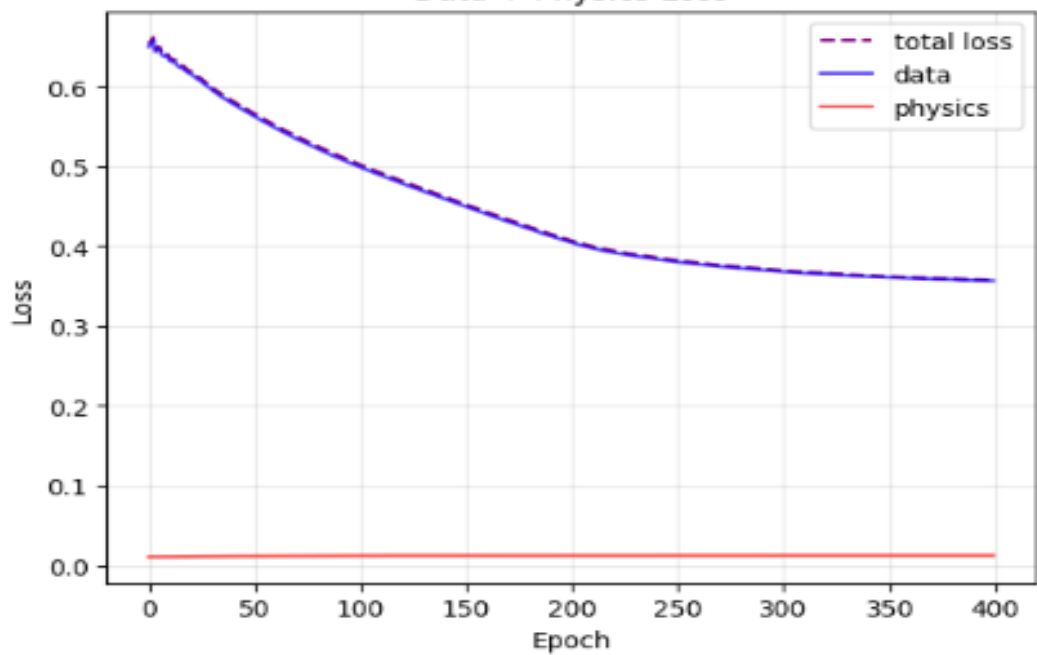


Figure 16: Data Loss+ Physics Loss Over Epochs

Parameter	PINN Value	ODE Value
( $\theta$ ) (supply air)	0.0018	0.0019
( $\theta$ ) (outdoor)	-0.0010	-0.0011
( $\theta$ ) (solar)	0.000038	0.000050

**Table 9: PINNs Learnable Parameters Value**

In this table, PINN and ODE values are together to check model successfully maintains the physical structure or not. Here, PINNs parameters closely matches the ODE results. In short, PINN model performs better than the FNN and ODE models in all testing scenarios. This close alignment between the differential equation parameters and the PINN-learned values further confirms the model accuracy also.

#### 4.7 XGBoost Model Results

XGBoost model was tested using the same input features as the other models. It was trained on one-step-ahead prediction and evaluated for 1-hour, 6-hours, and 12-hours prediction window. This arrangement allows for direct comparison with other models. However this parameters are selected based on trial and error not based on any optimization

Parameter	Value
Input features	Supply air, Outdoor temp, Solar radiation, Indoor temp
n_estimators	500
max_depth	4
learning_rate	0.05
subsample	0.8
reg_lambda (L2)	1.0
reg_alpha (L1)	0.1
Objective	reg:squarederror (MSE)
Scaler	RobustScaler

**Table 10: XGBoost Architecture Parameters Used In This Study**

One-Hour Model Graph:

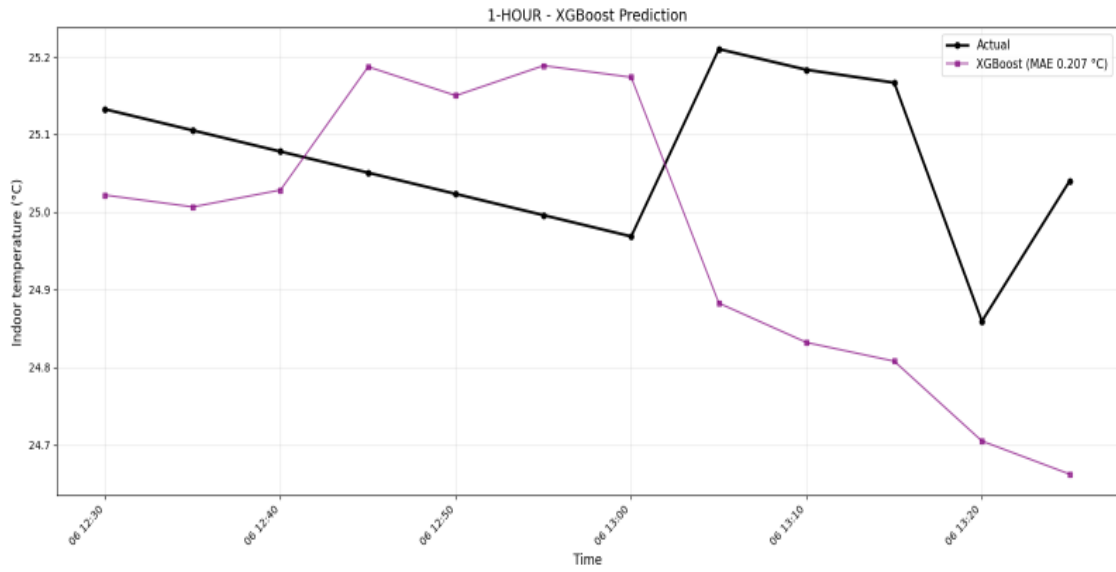


Figure 17: XGBoost One-Hour: Prediction vs Actual

Six-Hours Model Graph:

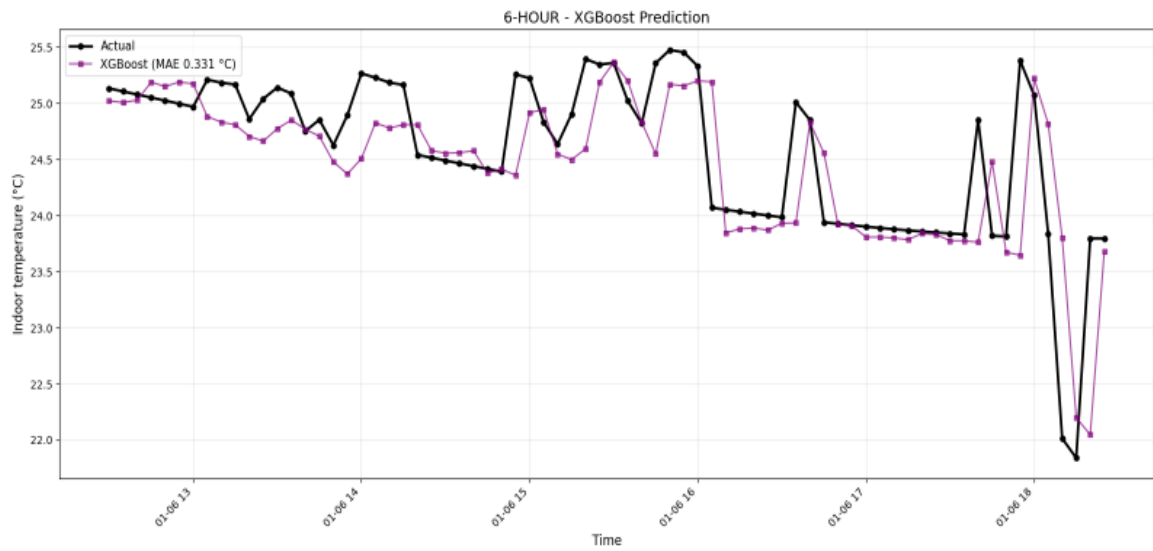
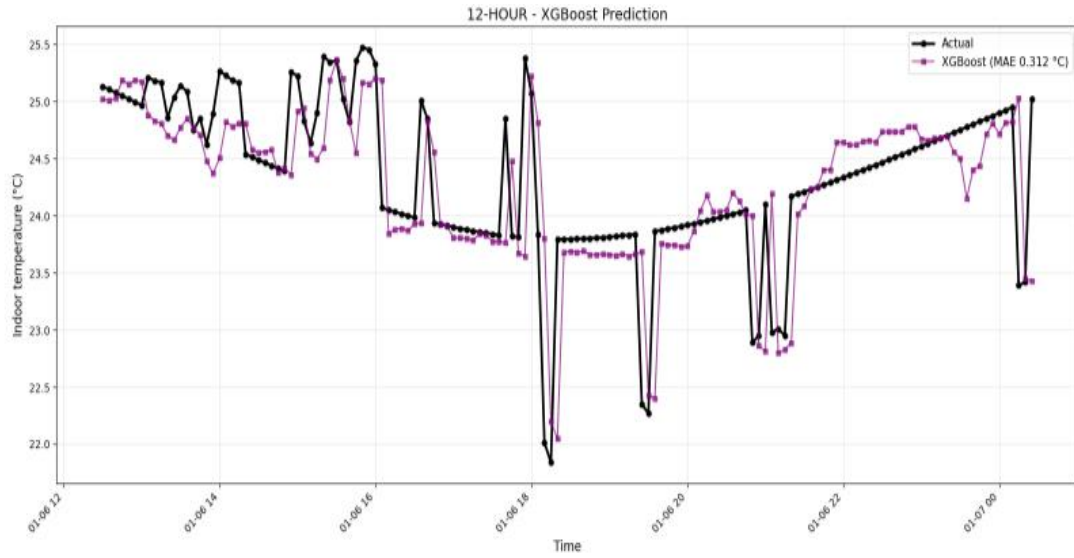


Figure 18 : XGBoost Six-Hours: Prediction vs Actual



**Figure 19: XGBoost Twelve-Hour : Prediction vs Actual**

Forecast Horizon	MAE (°C)	RMSE (°C)	R <sup>2</sup>
1-Hour	0.20	0.23	-1
6-Hours	0.33	0.52	0.45
12-Hours	0.31	0.50	0.46

**Table 11: XGBoost Prediction Performance**

XGBoost shows strong short-term predictive accuracy, It achieved the lowest 1-hour MAE among all models. Although, Performance gradually declines with longer prediction results. But still, XGBoost remains stable in MAE across 6 and 12 hours prediction windows and outperforms the FNN

## 4.8 Random Forest Model Performance

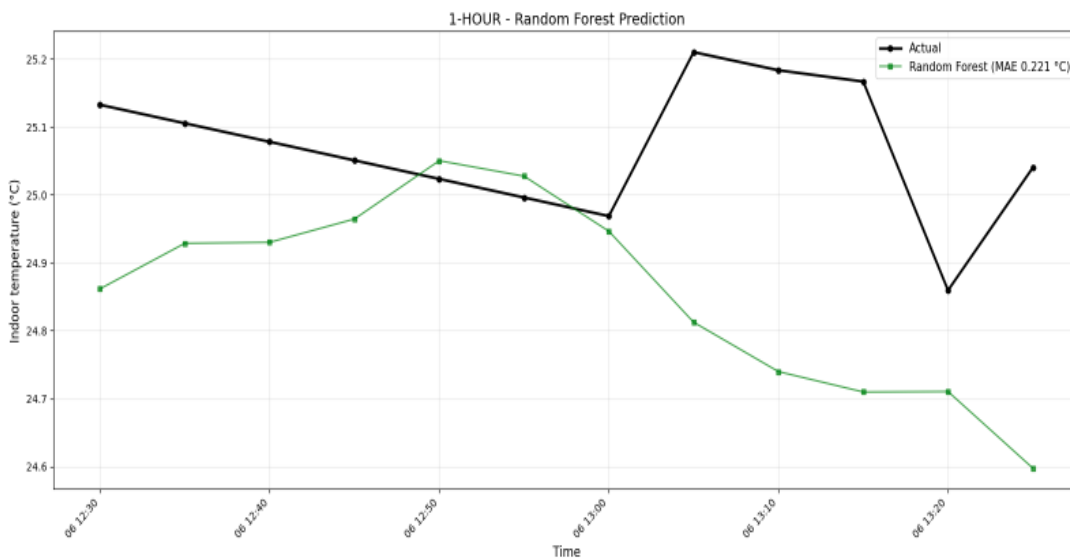
The RF regressor model was evaluated over three prediction time frames: 1- hour, 6- hours and 12 hours.

The model was trained using 15,801 and 5-minute samples and using same input features. and the hyperparameters were selected based on trial and error.

Parameter	Value
Input features	Supply air, Outdoor temp, Solar radiation, Indoor temp
Number of trees	500
Max depth	15
Min samples split	5
Min samples leaf	2
Max features	sqrt
Bootstrap	True
Loss metric	MSE (via internal splitting)
Scaler	RobustScaler
Random seed	42

**Table 12: Random Forest Parameters Used in This Study**

One-Hour Model Graph:



**Figure 20: One-Hour Forecast: Random Forest vs Actual**

Six-Hours Model Graph:

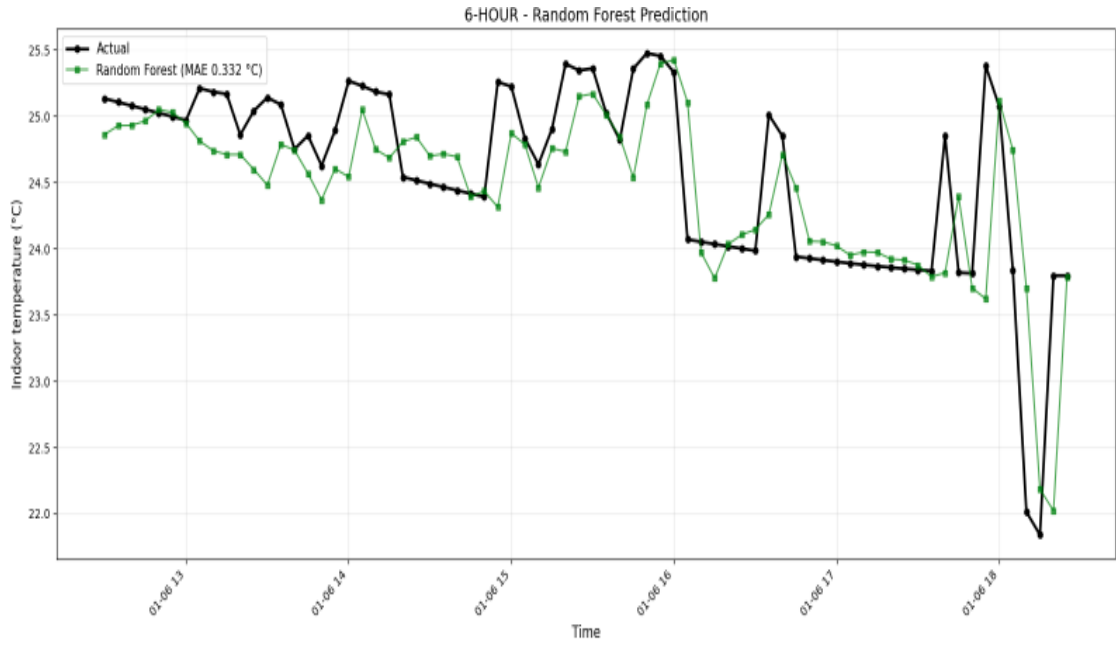


Figure 21: Six-Hour Forecast: Random Forest vs Actual

Twelve- Hours Model Graph:

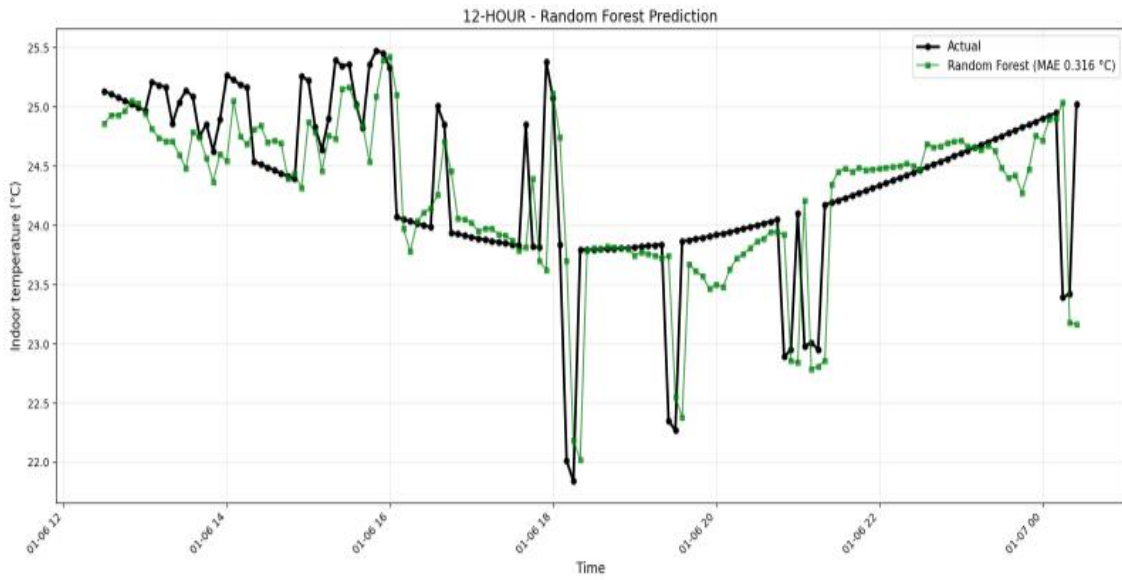


Figure 22: Twelve-Hours Prediction: Random Forest vs Actual

Horizon	MAE (°C)	RMSE (°C)	R <sup>2</sup>
1-Hour	0.22	0.27	-1
6-Hours	0.33	0.51	0.48
12-Hours	0.31	0.51	0.45

**Table 13: Random Forest Performance Prediction**

RF model shows strong short-term predictive performance among all the three horizons, achieving the lowest MAE of 0.221 °C in the 1-hour window. However, still for one hour horizon it shows highly unstable R<sup>2</sup> value. And across the 6-hours and 12-hour horizons RF model maintained consistent accuracy, achieving MAE values of 0.332 °C and 0.316 °C .

#### 4.9 LSTM Model Results

The LSTM network was performed for three forecast horizons: 1-hour, 6-hours, and 12-hours.

Parameter	Value
Input features	Supply air, Outdoor temp, Satellite value, Indoor temp
Sequence length	12 timesteps (1 hour)
Prediction horizon	1 step ahead (5 min), recursive
Architecture	LSTM(64) -> LSTM(32) -> Dense(16) -> Dense(1)
Activation functions	LSTM: tanh, Dense: ReLU
Optimizer	Adam (lr = 0.001)
Loss function	MSE
Batch size	32
Epochs	10 (early stopping)
Scaler	RobustScaler

**Table 14: LSTM Architecture Parameters Used In This Study**

## One-Hour Model Graph:

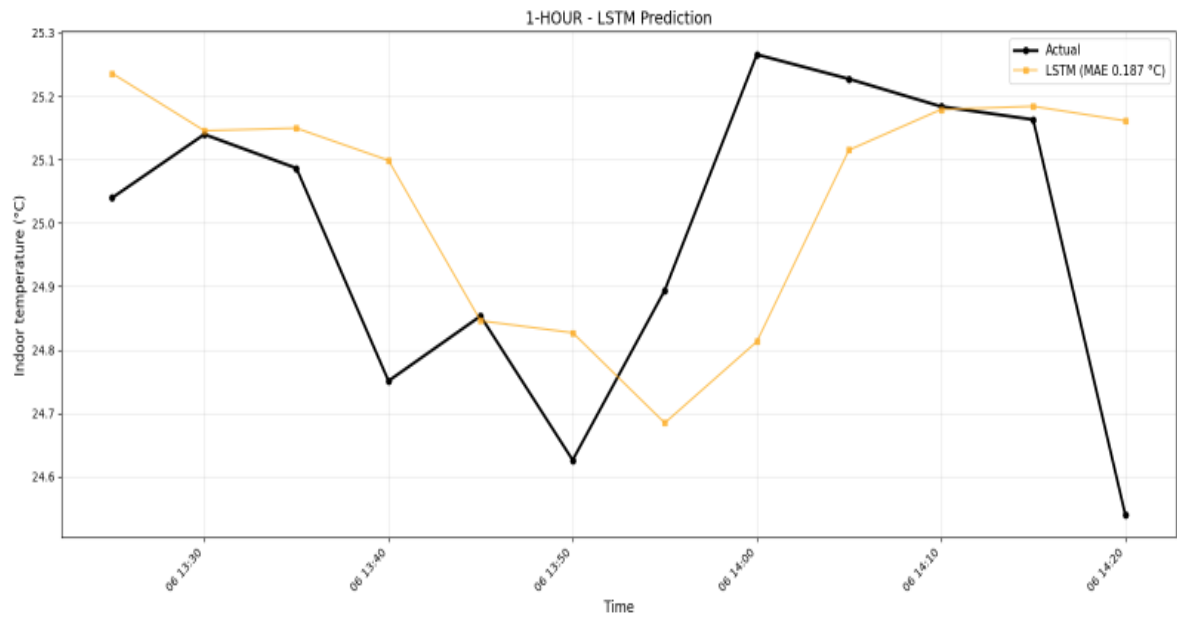


Figure 23: One-Hour LSTM Prediction vs Actual

## Six-Hours Model Graph:

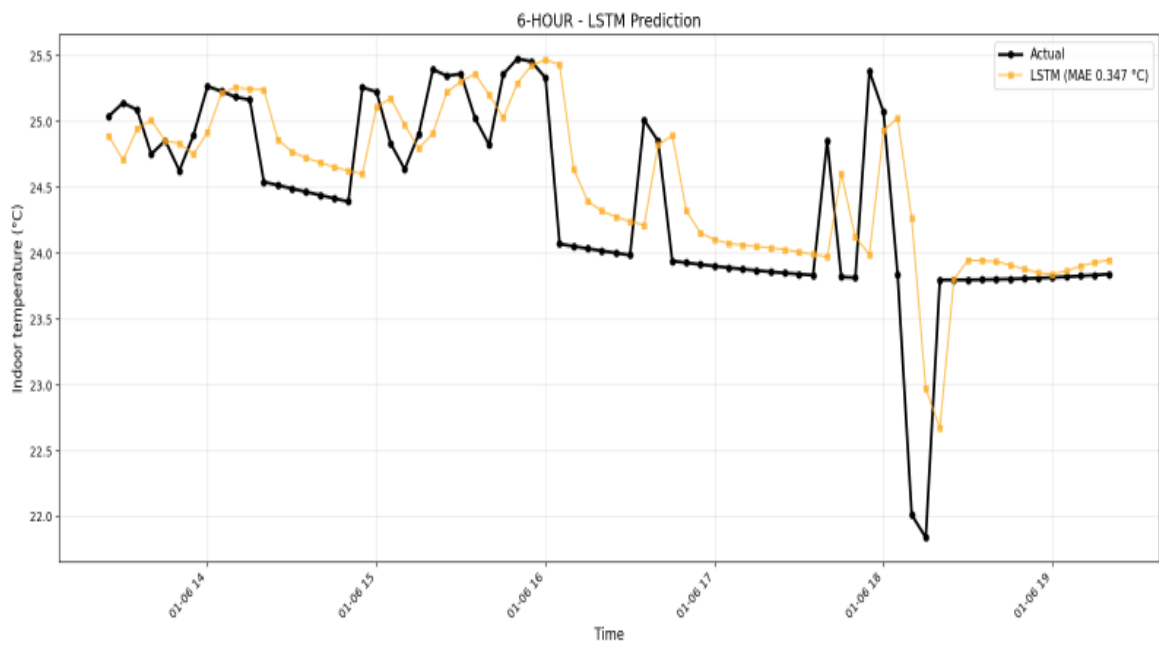


Figure 24: Six-Hour Prediction Plot

Twelve-Hours Model Graph:

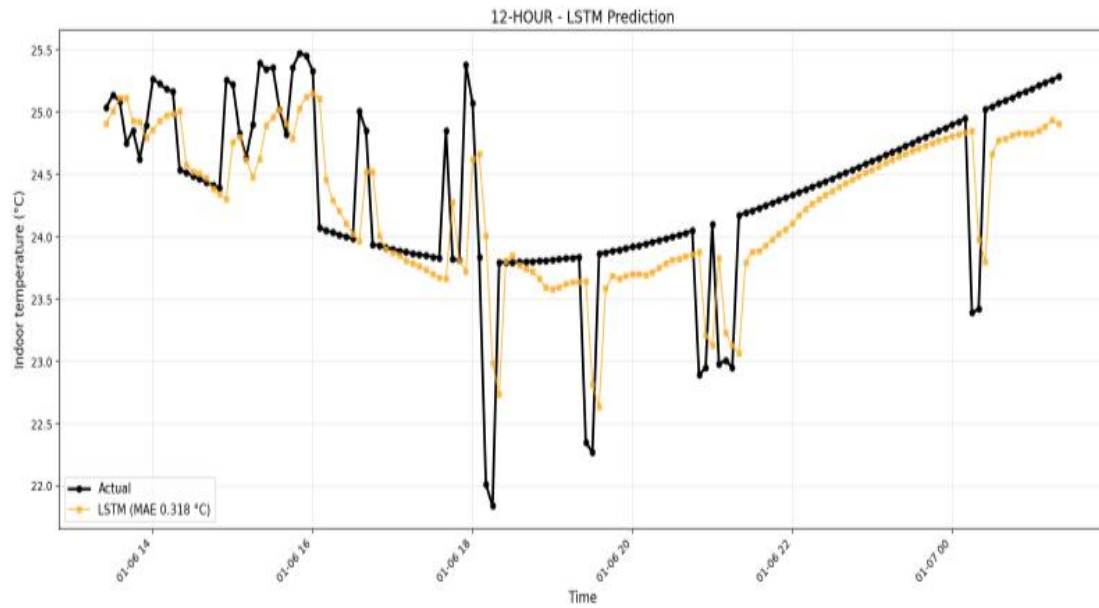


Figure 25: Twelve-Hour LSTM Prediction Plot

Horizon	MAE (°C)	RMSE (°C)	R <sup>2</sup>
1-Hour	0.18	0.26	-0.30
6-Hours	0.34	0.52	0.47
12-Hours	0.31	0.47	0.53

Table 15: LSTM Performance Prediction

#### 4.10 Model Performance Table

Model Name	Hours	MAE (°C)	RMSE (°C)	R <sup>2</sup>
<b>FNN</b>	1h	1.24	1.31	-1
	6h	0.60	0.62	-1
	12h	0.47	0.68	-0.29
<b>PINN</b>	1h	0.52	0.57	-1
	6h	0.36	0.34	0.34
	<b>12h</b>	<b>0.29</b>	<b>0.51</b>	<b>0.63</b>

<b>XGBoost</b>	1h	0.20	0.23	-1
	<b>6h</b>	<b>0.33</b>	<b>0.52</b>	<b>0.45</b>
	12h	0.31	0.50	0.46
<b>Random Forest</b>	1h	0.22	0.27	-1
	6h	0.33	0.51	0.48
	12h	0.31	0.51	0.45
<b>LSTM</b>	<b>1h</b>	<b>0.18</b>	<b>0.26</b>	<b>-0.30</b>
	6h	0.34	0.52	0.47
	12h	0.31	0.47	0.53

**Table 16: Combined Model Performance Table**

From the 12-hours prediction result, PINNs has the best performance, and it outperforms XGBoost by 6.4%, RF by 7.6%, LSTM by 8.1%, the FNN by 39.0% in terms of MAE.

From the 6-hours prediction result, XGBoost performed best, improving MAE by 4.3% over LSTM, 8.0% over the PINN, 45.1% over the FNN baseline. And it is performing nearly identically to RF.

From the 1-hour prediction result, LSTM achieved the lowest error, performing 9.9% better than XGBoost, 15.4% better than RF, 64.4% better than the PINN, and 84.9% better than the FNN model.

## 5 Conclusions

This thesis explored data-driven, physics-based, and hybrid physics-informed techniques for predicting indoor temperatures using building sensor and satellite data. Three categories of models were examined: ML models such as (FNN, RF, XGBoost, and LSTM), a ODE model, and the proposed PINN model, which incorporates the building's thermal dynamics directly into the learning process.

The results across the three forecasting horizons (1-hour, 6-hours, and 12-hours) show that no single model is universally optimal. Rather, performance depends on the prediction horizon and model structure. LSTM achieved the lowest error in the one-hour prediction, demonstrating the strength of sequence-based learning when dependencies are temporal. XGBoost performed best at the six-hours horizon, showing the effectiveness of tree-based models for medium-range forecasting. For the 12-hours prediction horizon, the physics-informed neural network outperformed all other models. It proves that Integrating physical constraints improved stability over longer periods.

Overall, this study indicates that integrating physical knowledge with data-driven learning can produce more precise and physically consistent forecasts, especially for longer prediction windows.

## 6 Limitations

Although in this research shows the potential of hybrid data-physics methods for indoor temperature prediction, but some limitations should be acknowledged.

First, The analysis was done using data from a single building and limited winter season period data. It is important to know that indoor temperature behavioural patterns is influenced by seasonal patterns, solar radiations, occupancy, ventilation, and heating strategies. It is possible that these conditions may vary across buildings. In this case missing sensor location information makes it impossible to assess what factors were cause of sudden changes. As sensor placement plays an important role in indoor environmental monitoring, and the lack of this information introduces uncertainty into both the training and evaluation of the models.

Second, in the dataset for indoor temperature data there are some sudden spikes. So it is recommended to install multiple sensors for reading. These sudden fluctuations can be sensor noise, calibration drift, or environmental disturbances. Since indoor temperature acts as the ground truth for this research model training and validation, such spikes reduce the reliability of performance metrics. So, In future additional sensors or redundant sensing could help smooth these anomalies and provide a more stable ground-truth reference.

Third, this dataset includes data from several more buildings but due to missing metadata and necessary sensors most of them are not feasible for research. As multiple buildings exist in the broader dataset, a significant portion of the buildings in the dataset lack critical variables. These variables include supply/return temperature, flow rate, humidity, and air ventilation. Many of the buildings also contain extensive amounts of irrelevant sensor data.

Finally, the indoor temperature is used as a ground truth contains sudden spikes not explained by the independent variables, making the prediction inaccurate. These models may perform better with proper ground truth data because then we can evaluate perfectly with the ground truth. Such improvements would better support for modelling and enable more robust research on prediction models.

## References

- Amasyali, K., & El-Gohary, N. M. (2018). A review of data-driven building energy consumption prediction studies. *Renewable and Sustainable Energy Reviews*, *81*, 1192–1205. <https://doi.org/10.1016/j.rser.2017.04.095>
- Du, H., Zhao, Z., Lyu, J., Li, J., Liu, Z., Li, X., Yang, Y., Lan, L., & Lian, Z. (2023). Gender differences in thermal comfort under coupled environmental factors. *Energy and Buildings*, *295*, 113345. <https://doi.org/10.1016/j.enbuild.2023.113345>
- Economidou, M., Todeschi, V., Bertoldi, P., D'Agostino, D., Zangheri, P., & Castellazzi, L. (2020). Review of 50 years of EU energy efficiency policies for buildings. *Energy and Buildings*, *225*, 110322. <https://doi.org/10.1016/j.enbuild.2020.110322>
- Häkkinen, A. (n.d.). *Machine learning models for predicting indoor temperature of central-heated residential buildings*.
- Hannula, E., Häkkinen, A., Solonen, A., Uribe, F., Wiljes, J. de, & Roininen, L. (2025). *Partially stochastic deep learning with uncertainty quantification for model predictive heating control* (No. arXiv:2504.03350). arXiv. <https://doi.org/10.48550/arXiv.2504.03350>
- Jaffal, I. (2023). Physics-informed machine learning for metamodeling thermal comfort in non-air-conditioned buildings. *Building Simulation*, *16*(2), 299–316. <https://doi.org/10.1007/s12273-022-0931-y>
- Jing, G., Ning, C., Qin, J., Ding, X., Duan, P., Liu, H., & Sang, H. (2023). Physics-guided framework of neural network for fast full-field temperature prediction of indoor environment. *Journal of Building Engineering*, *68*, 106054. <https://doi.org/10.1016/j.job.2023.106054>
- Li, D., Qi, Z., Zhou, Y., & Elchalakani, M. (2025a). Machine Learning Applications in Building Energy Systems: Review and Prospects. *Buildings*, *15*(4), 648. <https://doi.org/10.3390/buildings15040648>
- Li, L., Dai, S., Cao, Z., Hong, J., Jiang, S., & Yang, K. (2020). Using improved gradient-boosted decision tree algorithm based on Kalman filter (GBDT-KF) in time series

- prediction. *The Journal of Supercomputing*, 76(9), 6887–6900.  
<https://doi.org/10.1007/s11227-019-03130-y>
- Loffa, M. A., Macii, E., Patti, E., & Bottaccioli, L. (2025). Physics-Informed vs. Deep Learning: Indoor Temperature Prediction with Different Data Availability. *Proceedings of the 16th ACM International Conference on Future and Sustainable Energy Systems*, 742–750. <https://doi.org/10.1145/3679240.3734642>
- Lu, Y., Liu, C., Wang, K. I.-K., Huang, H., & Xu, X. (2020). Digital Twin-driven smart manufacturing: Connotation, reference model, applications and research issues. *Robotics and Computer-Integrated Manufacturing*, 61, 101837. <https://doi.org/10.1016/j.rcim.2019.101837>
- Luo, Z., Liu, X., Yang, Q., Qu, Z., Xu, H., & Xu, D. (2023). Numerical study on performance of porous brick roof using phase change material with night ventilation. *Energy and Buildings*, 286, 112972. <https://doi.org/10.1016/j.enbuild.2023.112972>
- Ma, N., Aviv, D., Guo, H., & Braham, W. W. (2021). Measuring the right factors: A review of variables and models for thermal comfort and indoor air quality. *Renewable and Sustainable Energy Reviews*, 135, 110436. <https://doi.org/10.1016/j.rser.2020.110436>
- Miller, C., Nagy, Z., & Schlueter, A. (2018). A review of unsupervised statistical learning and visual analytics techniques applied to performance analysis of non-residential buildings. *Renewable and Sustainable Energy Reviews*, 81, 1365–1377. <https://doi.org/10.1016/j.rser.2017.05.124>
- Muroni, A., Gaetani, I., Hoes, P.-J., & Hensen, J. L. M. (2019). Occupant behavior in identical residential buildings: A case study for occupancy profiles extraction and application to building performance simulation. *Building Simulation*, 12(6), 1047–1061. <https://doi.org/10.1007/s12273-019-0573-x>
- Park, J., Choi, H., Kim, D., & Kim, T. (2021). Development of novel PMV-based HVAC control strategies using a mean radiant temperature prediction model by machine learning in Kuwaiti climate. *Building and Environment*, 206, 108357. <https://doi.org/10.1016/j.buildenv.2021.108357>

- Raissi, M., Perdikaris, P., & Karniadakis, G. E. (2019). Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, *378*, 686–707. <https://doi.org/10.1016/j.jcp.2018.10.045>
- Sun, Y., Haghghat, F., & Fung, B. C. M. (2020). A review of the-state-of-the-art in data-driven approaches for building energy prediction. *Energy and Buildings*, *221*, 110022. <https://doi.org/10.1016/j.enbuild.2020.110022>
- Wang, Y., Yang, D., Yuan, Y., Zhang, J., & Au, F. T. K. (2025). Hybrid physics-informed neural network with parametric identification for modeling bridge temperature distribution. *Computer-Aided Civil and Infrastructure Engineering*, *40*(22), 3503–3524. <https://doi.org/10.1111/mice.13436>
- Yan, D., O'Brien, W., Hong, T., Feng, X., Burak Gunay, H., Tahmasebi, F., & Mahdavi, A. (2015). Occupant behavior modeling for building performance simulation: Current state and future challenges. *Energy and Buildings*, *107*, 264–278. <https://doi.org/10.1016/j.enbuild.2015.08.032>
- Zhao, J., Li, X., Shum, C., & McPhee, J. (2021). A Review of physics-based and data-driven models for real-time control of polymer electrolyte membrane fuel cells. *Energy and AI*, *6*, 100114. <https://doi.org/10.1016/j.egyai.2021.100114>

## Appendices

### Appendix 1. Github Link

<https://github.com/uddinminhaz/Integrating-Physics-Informed-and-Machine-Learning-Models-for-Indoor-Temperature-Prediction->